A

SEQUENCE LISTING

<110> University of Iowa Research Foundation et al.

<120> Novel polypeptides and methods of their use

<130> 875.006WO1

<150> US 60/105,575
<151> 1998-10-26

<160> 65

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 2265
<212> DNA
<213> Artificial Sequence

<220>
<223> A modified DNA sequence

<400> 1

```
gagggtagag aaatcgaatg ccccttgcat caaggtcggt ttgacgtttg cacaggcaaa      60
gccctgtgcg caccccgtgac acagaacatc aaaacatatc cagtcaagat tgagaacctg     120
cgcgtaatga ttgatttgag ctaagaattt taacaggagg cacccccgggc cctagagcgt     180
aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt      240
aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca      300
tgaactgaaa accatttttg cgcggaactg gcttttttctc actcatgata gcctgattcc     360
tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccggca      420
gaacgacggt tcgattcgtg cttttctgaa cgtttgccgg catcgtggca agacgctggt      480
gagcgtggaa gccggcaatg ccaaaggttt tgtttgcagc tatcacggct ggggcttcgg      540
ctccaacggt gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa      600
taaaaaatgt ctggggttga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta      660
cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggta      720
cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggt      780
tgtgatcaag gccaactgga aggcacccgc ggaaaacttt gtgggagatg cataccacgt      840
gggttggacg cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg      900
caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg      960
```

```
catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt    1020
gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc    1080
tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg    1140
ctcgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggaccta    1200
cgccattgtc gaaaaagaca tgcctgagga tctcaagcgc cgcttggccg actctgttca    1260
gcgaacggtc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc    1320
ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt    1380
cggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg    1440
cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cacgtcagca gctccaactg    1500
ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg    1560
ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac    1620
gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc    1680
actacgctgc tgacccagga agcgcatttg ttggacattc aggcttaccg tgcttggtta    1740
gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct    1800
tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa    1860
ctgaaagttc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg    1920
cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta    1980
cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc    2040
ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc    2100
cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg    2160
tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg    2220
aatgtacgtg tatgggcaat caacaagtcg tttcgataac cggtg                    2265
```

<210> 2
<211> 449
<212> PRT
<213> Artificial Sequence

<220>
<223> A polypeptide encoded by SEQ ID NO:1

<400> 2

```
Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
1               5                   10                  15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
            20                  25                  30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
        35                  40                  45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
    50                  55                  60
```

```
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65                  70                  75                  80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
        115                 120                 125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
145                 150                 155                 160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
    210                 215                 220
Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
                245                 250                 255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
            260                 265                 270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
        275                 280                 285
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
    290                 295                 300
Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305                 310                 315                 320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
                325                 330                 335
Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val
            340                 345                 350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
        355                 360                 365
Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
    370                 375                 380
```

```
Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
385                 390                 395                 400
Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
                405                 410                 415
Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
            420                 425                 430
Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
        435                 440                 445
Arg


<210> 3
<211> 9841
<212> DNA
<213> Artificial Sequence

<220>
<223> A modified DNA sequence

<400> 3
gaattcatca ggaagacatt caaatgaacg taaacaataa gggcagcgtc tgtatttgcg      60
gcagcgaaat gctccctaaa ttcctcattt accccatctg aggattgctt tatgacagta     120
aagtggattg aagcagtcgc tctttctgac atccttgaag gtgacgtcct cggcgtgact     180
gtcgagggca aggagctggc gctgtatgaa gttgaaggcg aaatctacgc taccgacaac     240
ctgtgcacgc atggttccgc ccgcatgagt gatggttatc tcgagggtag agaaatcgaa     300
tgccccttgc atcaaggtcg gtttgacgtt tgcacaggca aagccctgtg cgcacccgtg     360
acacagaaca tcaaaacata tccagtcaag attgagaacc tgcgcgtaat gattgatttg     420
agctaagaat tttaacagga ggcaccccgg gccctagagc gtaatcaccc ccattccatc     480
tttttaggt gaaaacatga attacaataa taaaatcttg gtaagtgaat ctggtctgag      540
ccaaaagcac ctgattcatg gcgatgaaga acttttccaa catgaactga aaaccatttt     600
tgcgcggaac tggctttttc tcactcatga tagcctgatt cctgcccccg gcgactatgt     660
taccgcaaaa atggggattg acgaggtcat cgtctcccgg cagaacgacg gttcgattcg     720
tgcttttctg aacgtttgcc ggcatcgtgg caagacgctg gtgagcgtgg aagccggcaa     780
tgccaaaggt tttgtttgca gctatcacgg ctggggcttc ggctccaacg gtgaactgca     840
gagcgttcca tttgaaaaag atctgtacgg cgagtcgctc aataaaaaat gtctggggtt     900
gaaagaagtc gctcgcgtgg agagcttcca tggcttcatc tacggttgct tcgaccagga     960
ggcccctcct cttatggact atctgggtga cgctgcttgg tacctggaac ctatgttcaa    1020
gcattccggc ggtttagaac tggtcggtcc tccaggcaag gttgtgatca aggccaactg    1080
gaaggcaccc gcggaaaact ttgtgggaga tgcataccac gtgggttgga cgcacgcgtc    1140
ttcgcttcgc tcgggggagt ctatcttctc gtcgctcgct ggcaatgcgg cgctaccacc    1200
```

```
tgaaggcgca ggcttgcaaa tgacctccaa atacggcagc ggcatgggtg tgttgtggga    1260
cggatattca ggtgtgcata gcgcagactt ggttccggaa ttgatggcat tcggaggcgc    1320
aaagcaggaa aggctgaaca aagaaattgg cgatgttcgc gctcggattt atcgcagcca    1380
cctcaactgc accgttttcc cgaacaacag catgctgacc tgctcgggtg ttttcaaagt    1440
atggaacccg atcgacgcaa acaccaccga ggtctggacc tacgccattg tcgaaaaaga    1500
catgcctgag gatctcaagc gccgcttggc cgactctgtt cagcgaacgg tcgggcctgc    1560
tggcttctgg gaaagcgacg acaatgacaa tatggaaaca gcttcgcaaa acggcaagaa    1620
atatcaatca agagatagtg atctgctttc aaaccttggt ttcggtgagg acgtatacgg    1680
cgacgcggtc tatccaggcg tcgtcggcaa atcggcgatc ggcgagacca gttatcgtgg    1740
tttctaccgg gcttaccagg cacacgtcag cagctccaac tgggctgagt tcgagcatgc    1800
ctctagtact tggcatactg aacttacgaa gactactgat cgctaacaga cgagtcgacc    1860
atgatgatca atattcaaga agacaagctg gtttccgccc acgacgccga agagattctt    1920
cgtttcttca attgccacga ctctgctttg caacaagaag ccactacgct gctgacccag    1980
gaagcgcatt tgttggacat tcaggcttac cgtgcttggt tagagcactg cgtggggtca    2040
gaggtgcaat atcaggtcat ttcacgcgaa ctgcgcgcag cttcagagcg tcgttataag    2100
ctcaatgaag ccatgaacgt ttacaacgaa aattttcagc aactgaaagt tcgagttgag    2160
catcaactgg atccgcaaaa ctggggcaac agcccgaagc tgcgctttac tcgctttatc    2220
accaacgtcc aggccgcaat ggacgtaaat gacaaagagc tacttcacat ccgctccaac    2280
gtcattctgc accgggcacg acgtggcaat caggtcgatg tcttctacgc cgcccgggaa    2340
gataaatgga aacgtggcga aggtggagta cgaaaattgg tccagcgatt cgtcgattac    2400
ccagagcgca tacttcagac gcacaatctg atggtctttc tgtgattcag tgaccatttt    2460
tacaaatggt cactgcaacc gcggtcacca ttaatcaaag ggaatgtacg tgtatgggca    2520
atcaacaagt cgtttcgata accggtgcag gctcaggaat cggtctcgaa ctggttcggt    2580
cctttaagtc ggccggttat tacgtatccg ctctcgtacg aaacgaggag caagaggcgc    2640
ttctttgcaa agagttcaag gacgcactcg agattgtagt gggcgatgtc cgggaccacg    2700
caacaaatga gaagctgata aagcaaacaa tcgatagatt cggtcatctt gattgtttta    2760
ttgcaaatgc cggtatctgg gattacatgc tgagcatcga agagccttgg gagaaaatat    2820
cgagcagttt tgacgaaata ttcgacatta atgtcaagag ctatttcagt ggcatcagtg    2880
ccgccctgcc ggaactgaaa aagactaacg gatcagtggt gatgaccgct tcggtgtcgt    2940
cccatgcggt cggtggtggt ggttcttgct acatcgccag caagcatgcg gtgctcggta    3000
tggttaaggc tttggcctac gaattggccc ccgaagttcg cgtgaacgct gtttcgccgg    3060
ggggcaccgt gacgtctctg tgcggtcccg cgagcgccgg tttcgacaaa atgcacatga    3120
aagacatgcc cggcatcgac gatatgatca aaggtctcac gcctcttggg tttgcagcca    3180
agcccgaaga cgtggtggca ccctatttgt tgctggcttc gcgaaagcaa ggaaaattca    3240
tcaccggcac cgtgattagc attgatggcg gtatggcgct cggtcgcaag tgagcttgta    3300
gccgatcaga agttatagac acatttcagg tgacgcccca tgaagacaaa actgtttatc    3360
aataacgcct ggatcgattc tagtgaccag cagaccttcg agcgcataca ccccgtcagc    3420
agcgatgtgg tgactgagag cgcaaacgcc acagtgacgg acgcgataaa ggcggcgcaa    3480
gcggccgagg aggcgttcaa gacctggaag gccgttggac cttcagagcg tcgccgcctt    3540
ctcctaaagg tcgccgatgt catggaaagt aaaacaccca agttcatcga agtgatggcc    3600
```

```
atggaggtgg gagcttccgc cctttgggcc ggattcaacg tccatgcgtc tgccaatgtg    3660
ttccgagagg ctgcctcgct ggctacccaa attcagggtg aaaccatccc aacggacaaa    3720
gccgaaacgc tctcaatgac actacgtcag ccggtcggcc cgatcctaag catcgttcca    3780
tggaacggca ccgcagtgct tgcggcacga gccatcgctt atccgctggt ctgtggcaac    3840
actgtggtgt tcaaaggctc tgaatttagt cccgcgacgc atgccctgat cacccagtgc    3900
gtgcaggaag ccgggctgcc cgctggcgtg ctcaattacc tcaactcttc gcctgaccgt    3960
tcgcccgaga tcgctgacgc actgatctct gccaaggaga tccgccgcat caacttcacg    4020
ggttccaccc gcgtgggcag cattatcgcg cagaaagccg cgcaacacct caagcgctgc    4080
ctgctggagc tcggcggcaa gtccccgctt attgttctgg atgatgcaga catcgatgcg    4140
gcggtcaagg cagcggtgtt cggtagcttc ctgttccaag gtcagatctg catgtccact    4200
gagcgcttga tcgttgatga gaagatagcc gacgaatttg tcgcaaaatt tgtcgaaaaa    4260
actaagcgct tgagcgcagg cgacccgtgc gtaactggcg actgcatcat cggcccgatg    4320
gtctcgccaa attcgggtga gcggatcaat ggtttgttca aagacgcgat cgacaaaggg    4380
gcaaaagttg tttgcggcgg cttggcccaa ggtgcgctca tgccggccac gatcctggat    4440
cacgtcaaat ctgacatgcg gatttacgat gaggagacct ttggtcccat caccgtggta    4500
atccgttgta aaggcgaagc agaggccgtc cgcattgcca acgacagcgt ctatggcctg    4560
tcgtcgggcg tatttggccg cgacatcaac cgcgctctac gcgtgggtat gtccatcgaa    4620
tatggttctg tacacatcaa cggttcgacc gtccagaacg aggcgcaggc tccttacgga    4680
ggcaccaaga acaccggcta cgggcgcttc gacggccgtg ctgtaatcga cgagttcaca    4740
gagatcaagt ggctgaccat cgaacctttc gagcagcaat atcccttctg ataagcacta    4800
actcccagga atcaaactat gagtaagcaa gctgcagtta tcgagctcgg atacatgggt    4860
atctcggtca aggaccctga tgcgtggaaa tcatttgcca cggatatgct aggtctgcaa    4920
gttcttgatg agggtgagaa ggaccgtttc tatctgcgga tggattactg gcatcatcgg    4980
atcgtagtcc atcacaacgg acaggacgac ttggagtacc taggctggcg tgtagccggc    5040
aagccggagt tcgaagctct gggtcaaaag cttattgatg ccggttacaa gatccgcatc    5100
tgcgacaaag ttgaggctca ggagcgtatg gtgttgggtc tgatgaagac agaagatccg    5160
ggcggcaacc cgaccgagat attctggggc ccccggatcg acatgagcaa cccgttccat    5220
cccggtcgcc ccctgcacgg aaagtttgtg accggtgacc aaggcttggg ccattgcatc    5280
gttcgccaaa ccgacgtcgc agaagctcat aagttttata gcctgctggg cttccgtggg    5340
gacgtcgaat accggattcc gttgcccaac ggcatgactg ccgaactgtc gttcatgcat    5400
tgcaacgccc gtgatcactc cattgctttt ggtgccatgc ccgctgccaa acgactcaat    5460
cacttgatgc ttgagtacac ccatatggaa gacttgggat acacgcacca acagtttgta    5520
aagaacgaaa ttgacattgc cttgcagctt ggcattcacg ccaacgacaa ggcgttgacg    5580
ttctatggtg caacgccttc gggctggctc attgagcccg gctggcgagg tgccacggcc    5640
atagatgaag cggagtatta cgtcggcgac atcttcggcc atggcgtgga ggccactgga    5700
tatggcctgg atgtaaaact gagctaaaga tgcgcgctcg ttgggcgagg ctctagtcca    5760
gcatcttcat acgcaaccaa ccttgcaggg cgatgagatc aaaggacgtt aaagcgaagg    5820
ggaagtggtt cgggccatgc gcataccgat ccatgacatt tgtttcatag tatataggta    5880
gataggtgaa tcaagcgctt agtcaactag tggacacatc tgttccatga ggctatctac    5940
tatctattca aaacaagaat aataaatagg atgaaaataa taatgataaa aagaacgatt    6000
```

```
tgtcttgtgt atcctctatt ctgtttggca agccccacat gggccgaaga gtcgccttgg    6060
acgtaccgta ttggtatgac taatgtagct ttcgatgcta gcgcaaaagt atacttaaat    6120
ggtcagcggg tgccaggagg aagcgctgat gcgagcgata acaacgcgct tacattcgac    6180
ttcggctacg ccatcaacga ccagtggaat gtacgtgcga ttgtcggtat tccgcctaca    6240
actaaagtga cgggcgcagg cacacttcct ggtatccagc tggggaaaat aacttacgct    6300
ccaacagtat taacgttgaa ctataacctc cccgctttgg gtcccgttcg ccctcacata    6360
ggtgcgggag tcaattacac gcggattttt gaaagtcggg acgctaatct aaaatcgttc    6420
gatgccgacc acgcttggtc ccccgcgcta catgttggtg ccgatattga cgttaaccgt    6480
ggttggttcg ttagcattga tatccggaag ttatacctga aaaccgacgc atcagggtac    6540
ttggggccac aggaggctaa agcacgggta actcttgacc cattactaac ttcgatcgcg    6600
atcggacgcc aattctgatg attctgttta aagttcttta tctatctaac cgcaaagggt    6660
gtttccatgt cgaataaaat tatgaaaacg tcgcgtctta ccgccgaaga tatcaacggc    6720
gcctggacta taatgcccac accctcgacg cctgatgctt ctgattggcg cagcactgcc    6780
accgtggact tagaagagac tgcccgcata gttgaagagc tgattgcagc tggtgtcaac    6840
ggtattctaa gtatgggtac ttttggtgag tgcgccacgt tgacctggga tgaaaaacgt    6900
gattatgtct cgacgattgt cgagaccatt cgtggtcgcg tgccttattt ctgtggcacg    6960
acagccttaa atacccgaga agtcatccgc cagacccgag agcttatcga tattggcgcc    7020
aacggcacta tgctcggggt gccgatgtgg gtgaagatgg acctgcctac agcggttcag    7080
ttctatcgtg atgttgcaga tgcggtacca gaggctgcca ttgcgattta cgccaacccc    7140
gaagcattca agttcgactt ccctcgccca ttctgggcag agatgtccaa aattccgcag    7200
gtagtgactg cgaagtatct aggcatcgga atgcttgact tggacctgag actggcaccc    7260
aacatccgct tccttcccca cgaagatgac tattacgcgg ccgcacgcat caatcccgag    7320
cgcataaccg cgttctggtc aagcggggcc atgtgcggcc cggctaccgc catcatgttg    7380
cgtgacgaag tggtgcgggc caagagcacc ggtgactggg ccaaggccaa agccatctcc    7440
gatgatatgc gtgcagccga ctcgacattg tttccgcgtg gcgacttttc ggagttctcg    7500
aagtataata tcgggcttga aaaggcacgg atggatgcgg ctggttggct caaggctggg    7560
ccctgccgtc cgccctacaa ccttgttcca gaagactacc tcgctggtgc acagaaatca    7620
ggcaaggctt gggccgcgct gcacgctaaa tacagtaatg aattgaagta gttcacctcc    7680
gcagacctga gtgacagggt ggcgcagacg ctgagggtgc aggaattaag tgagctaaag    7740
cacatttctt gcgccaggca ttgccagatc agcaaagttt gctgatctgg cagtttcaaa    7800
aatttgggcg aaagctgata tcaggaatac gggataaagg cagtgcacca taacgacggg    7860
gcgtgccatt cgtgatgaac gattttgcta ttgtgccgac ttctgttctt ggagtgtttg    7920
attgtgattg tcgattttta tttcgatttt ttgagtccgt tctcttactt ggccaaccat    7980
cgtttgtcaa agcttgcgca agactatggc ttttccattc gttattacgc aatcgatttg    8040
gcgcgagtta aaatagccat cggaaacgtt ggtccatcta atcgcgacct gatagtcaag    8100
ctggactatt tgaaagtaga tttgcaacgg tgggccgagc tttacgaaat accgttggta    8160
ttcccagcta actacaacag ccgacggatg aatactgggc tttattactc gggagccatg    8220
gcacagactg gtgcctatgt gaatgtagta tttaatgcgg tttggggaga tggcatagct    8280
ccagatttgg aaagcttgcc tgctctggta tctgaaaaac taggctggga tcgtagcgcc    8340
ttcgaggact ttatcagcag cgatgccgca acagagaggt atgacgagca gacacatgcc    8400
```

```
gcgatcgaac gcaaagtgtt cggtgtgcca acgatgtttt tgggcgatga aatgtggtgg    8460
ggaaacgacc gtctatttat gctcgagaac gcagtgggag gtgcgcctgt aaatggagaa    8520
tagtcgctac ggagcgcttg tgccggctaa atgccgatat aagtggttga cctgatcgtt    8580
atttgctcga tacagcgctt tcaaaatcag cggctactga agtcagataa aaatgcggga    8640
ctacttcagg catcctgtgc gacacaaagt tttacctgta attgtccacc tattccgagt    8700
ttggaatggt agctgactcg ctatgcgacc agcgatagcc taacaagaca tgcatcactg    8760
gtaacggtgg ggtgtgaagc tcctgcaaca atgtagcccc ttgatgtgtg tatttgctgc    8820
gaggtgaagc acagatgctc ggagccgtac cggcttgtgg cgctaggctg gcaagtatga    8880
gcaacgtaag tgggggttgg ggcgcaatgg gaaccaaaaa ccaacgcaag ccttaccagc    8940
gtcgttcggt gccttcctcc catgcctccg cctcgataaa gcagctgcgc atatcggctt    9000
cctggctgat ctcggttagt aggtcatgca aggtcttgtc cagcgcctcg tcgctccgat    9060
acggaatggt cagctcgtaa tggccggtct ccgaccgctt catgccgtag ggctccaggc    9120
agtagcgctc gatgttctcc gtggcccgct tccgaccgcg catgaacttg ctgttgttca    9180
ccaccgccag gcgcagggtg acggtggcca cccgctcgac ggttgactct gccggtgacg    9240
cgatattgcg cttttgacct cgcgccaggg cgctcttctg gtacgtcccg atctcgacgc    9300
cacggtggcg taggtagctg tacagggtgc tcttggagat gtgcaacttc tcgccgatgg    9360
cgctgacgct caggcggccc tcgcggtaca gggtctccgc cgccatggcg gtggcctcgg    9420
ccttggctgg caggcccttg ggacggcgac cgatccggcc tcgagtccgt gccgccgaca    9480
ggcccgcctg agtccgctcg cggatcagct cgcgctcgaa ctgaacaggt tgaacaccag    9540
gcgatcttgg gcgtgggtgc tgtcaatggg gtcgttcagg ctctgcaagc cgactttgcg    9600
tgcagccagc tagccgacca actcaaccag gtgcttgagc gagcgaccga ggcgatccag    9660
cttccagatc accacggcat cgcccgctcg aacatgggct agcaacttgt ccaactccgg    9720
ccgcgcgctt tttgcgccgc tggcgatgtc ttgatagatg cgttcgcacc cggcctgttt    9780
cagggcatcg acctggaggt cggcgttcta atcccgagtg ctcacccgcg tataaccgat    9840
c                                                                    9841
```

```
<210> 4
<211> 2515
<212> DNA
<213> Artificial Sequence

<220>
<223> A modified DNA sequence

<400> 4
gaattcatca ggaagacatt caaatgaacg taaacaataa gggcagcgtc tgtatttgcg      60
gcagcgaaat gctccctaaa ttcctcattt accccatctg aggattgctt tatgacagta     120
aagtggattg aagcagtcgc tctttctgac atccttgaag gtgacgtcct cggcgtgact     180
```

```
gtcgagggca aggagctggc gctgtatgaa gttgaaggcg aaatctacgc taccgacaac      240
ctgtgcacgc atggttccgc ccgcatgagt gatggttatc tcgagggtag agaaatcgaa      300
tgccccttgc atcaaggtcg gtttgacgtt tgcacaggca aagccctgtg cgcacccgtg      360
acacagaaca tcaaaacata tccagtcaag attgagaacc tgcgcgtaat gattgatttg      420
agctaagaat tttaacagga ggcaccccgg gccctagagc gtaatcaccc ccattccatc      480
tttttaggt gaaaacatga attacaataa taaaatcttg gtaagtgaat ctggtctgag       540
ccaaaagcac ctgattcatg gcgatgaaga acttttccaa catgaactga aaaccatttt      600
tgcgcggaac tggcttttc tcactcatga tagcctgatt cctgcccccg gcgactatgt       660
taccgcaaaa atggggattg acgaggtcat cgtctcccgg cagaacgacg gttcgattcg      720
tgcttttctg aacgtttgcc ggcatcgtgg caagacgctg gtgagcgtgg aagccggcaa      780
tgccaaaggt tttgtttgca gctatcacgg ctggggcttc ggctccaacg gtgaactgca      840
gagcgttcca tttgaaaaag atctgtacgg cgagtcgctc aataaaaaat gtctggggtt      900
gaaagaagtc gctcgcgtgg agagcttcca tggcttcatc tacggttgct tcgaccagga      960
ggcccctcct cttatggact atctgggtga cgctgcttgg tacctggaac ctatgttcaa     1020
gcattccggc ggtttagaac tggtcggtcc tccaggcaag gttgtgatca aggccaactg     1080
gaaggcaccc gcggaaaact ttgtgggaga tgcataccac gtgggttgga cgcacgcgtc     1140
ttcgcttcgc tcgggggagt ctatcttctc gtcgctcgct ggcaatgcgg cgctaccacc     1200
tgaaggcgca ggcttgcaaa tgacctccaa atacggcagc ggcatgggtg tgttgtggga     1260
cggatattca ggtgtgcata gcgcagactt ggttccggaa ttgatggcat tcggaggcgc     1320
aaagcaggaa aggctgaaca aagaaattgg cgatgttcgc gctcggattt atcgcagcca     1380
cctcaactgc accgttttcc cgaacaacag catgctgacc tgctcgggtg ttttcaaagt     1440
atggaacccg atcgacgcaa acaccaccga ggtctggacc tacgccattg tcgaaaaaga     1500
catgcctgag gatctcaagc gccgcttggc cgactctgtt cagcgaacgg tcgggcctgc     1560
tggcttctgg gaaagcgacg acaatgacaa tatggaaaca gcttcgcaaa acggcaagaa     1620
atatcaatca agagatagtg atctgctttc aaaccttggt ttcggtgagg acgtatacgg     1680
cgacgcggtc tatccaggcg tcgtcggcaa atcggcgatc ggcgagacca gttatcgtgg     1740
tttctaccgg gcttaccagg cacacgtcag cagctccaac tgggctgagt tcgagcatgc     1800
ctctagtact tggcatactg aacttacgaa gactactgat cgctaacaga cgagtcgacc     1860
atgatgatca atattcaaga agacaagctg gtttccgccc acgacgccga agagattctt     1920
cgtttcttca attgccacga ctctgctttg caacaagaag ccactacgct gctgacccag     1980
gaagcgcatt tgttggacat tcaggcttac cgtgcttggt tagagcactg cgtggggtca     2040
gaggtgcaat atcaggtcat ttcacgcgaa ctgcgcgcag cttcagagcg tcgttataag     2100
ctcaatgaag ccatgaacgt ttacaacgaa aattttcagc aactgaaagt tcgagttgag     2160
catcaactgg atccgcaaaa ctggggcaac agcccgaagc tgcgctttac tcgctttatc     2220
accaacgtcc aggccgcaat ggacgtaaat gacaaagagc tacttcacat ccgctccaac     2280
gtcattctgc accgggcacg acgtggcaat caggtcgatg tcttctacgc cgcccgggaa     2340
gataaatgga aacgtggcga aggtggagta cgaaaattgg tccagcgatt cgtcgattac     2400
ccagagcgca tacttcagac gcacaatctg atggtctttc tgtgattcag tgaccatttt     2460
tacaaatggt cactgcaacc gcggtcacca ttaatcaaag ggaatgtacg tgtat          2515
```

<210> 5

<211> 9706.

<212> DNA

<213> Artificial Sequence

<220>

<223> A modified DNA sequence

<400> 5

```
gctgatttcg accaattatt ccacgccgtc tgcgccagca aggccgtaat cagttcgtaa      60
aggggagccc gcagtaacag cgaaccgacg ccggtagctg cgcgcgtgaa tcctgagatc     120
acaaattctt ggcgattatc gcatttcaat ttcacccaag cgccgttcct tgcgctgctg     180
tagcggcggc catcacgctt ctccacgatt ccctccagct gcatccggca agcgttggcc     240
agcagctgca ctgggtaagc ttccagggta tcgctgaagc gtacgcgctc caatgggcag     300
tgctcgagga gcgcctcaag tacctggcgg ctcagttcaa cgccctctcc ccgcagatct     360
gcaccgtcga ggtagagcag gtcgaacgcc acgtacacaa atttgtcggt acaccgtgcg     420
gcgaacgctg actgcagggc ttgaaatgct ggtcgaccgt cgtcatactg gaatacgacc     480
tcaccgtcga gccaggctga atggacggct aggcgtccca gctatttggc caggagtggc     540
atatgatcga tctagtcaaa cccgtttttg gtgaaaagct gcacctggtc accatcgatc     600
cgtgctaaca gacgatacca gtcgtacttg atctcgtatc gccatgctcc tgccgacgac     660
agtgtcagta gagtaaccag ttgtggcgcg atccaacacg ggtttggtgc cttgggcggt     720
ttcactatcg ccacctcgct ataggaacat tccttcaggg tggagtgcgt aattttctga     780
aaggggagcc aggttatgag tattcacatt ggtgataaac aacatcactt atgcgttatt     840
gacatataac gtcgtattca cgattattta ccatataagt cttataataa cgaagccata     900
ttatggaact cctcatacaa ccgaaaaatc gcataattcc cttcagtgcc ggtgccaacc     960
ttctggaagt gcttcgcgag aacggtgtag ctatttccta cagttgcttg tctgggcgtt    1020
gcggaacctg tcgctgccgg gttatagatg gcagtgtcat tgattctggg gcggaaaatg    1080
ggcaatcaaa cctcaccgac aagcagtatg tgctcgcctg tcagtcagta ctcactggca    1140
attgcgctat cgaagtccca gaagccgacg aaattgtcac tcacccggcg cgaatcatca    1200
agggcacagt ggtcgcagtc gagtcgccca ctcacgatat ccgtcgctta cgcgtacgcc    1260
tctccaagcc cttcgagttc tcacccggac agtacgcgac actgcagttc agccctgagc    1320
atgcgcgtcc gtattcaatg gcaggtttgc cagatgacca agaaatggag ttccacatac    1380
gcaaggtgcc gggtgggcgc gtcacggagt atgttttcga acacgtccgc gaaggtacaa    1440
gcatcaagtt gagcgggcct cttggtacgg cttatctacg tcagaagcac accggaccga    1500
tgctgtgtgt aggtggcggg accggactcg caccggtgct gtcgattgtt cgcggcgcgc    1560
tgaagtcggg tatgacgaac cccatcctcc tttatttcgg ggtgcgcagt cagcaagacc    1620
tctacgacgc agagcgattg cacaaactcg ccgctgacca ccctcaactg accgtacaca    1680
cggtgattgc aacgggcccg attaatgagg gtcagcgagc cggcctaatt accgatgtga    1740
tcgaaaaaga catcctttcg ctggctgggt ggagggccta cctgtgcggc gcaccagcga    1800
tggttgaagc gttgtgcacc gtcaccaagc atcttggaat atcacccgaa catatttatg    1860
```

```
ccgatgcctt ctatcccggt gggatctgaa tagttcccgg ccatgcacct ctgtccatcg    1920
agaattcatc aggaagacat tcaaatgaac gtaaacaata agggcagcgt ctgtatttgc    1980
ggcagcgaaa tgctccctaa attcctcatt taccccatct gaggattgct ttatgacagt    2040
aaagtggatt gaagcagtcg ctctttctga catccttgaa ggtgacgtcc tcggcgtgac    2100
tgtcgagggc aaggagctgg cgctgtatga agttgaaggc gaaatctacg ctaccgacaa    2160
cctgtgcacg catggttccg cccgcatgag tgatggttat ctcgagggta gagaaatcga    2220
atgccccttg catcaaggtc ggtttgacgt ttgcacaggc aaagccctgt gcgcacccgt    2280
gacacagaac atcaaaacat atccagtcaa gattgagaac ctgcgcgtaa tgattgattt    2340
gagctaagaa ttttaacagg aggcaccccg ggccctagag cgtaatcacc cccattccat    2400
cttttttagg tgaaaacatg aattacaata ataaaatctt ggtaagtgaa tctggtctga    2460
gccaaaagca cctgattcat ggcgatgaag aacttttcca acatgaactg aaaaccattt    2520
ttgcgcggaa ctggcttttt ctcactcatg atagcctgat tcctgccccc ggcgactatg    2580
ttaccgcaaa aatggggatt gacgaggtca tcgtctcccg gcagaacgac ggttcgattc    2640
gtgcttttct gaacgtttgc cggcatcgtg gcaagacgct ggtgagcgtg gaagccggca    2700
atgccaaagg ttttgtttgc agctatcacg gctggggctt cggctccaac ggtgaactgc    2760
agagcgttcc atttgaaaaa gatctgtacg gcgagtcgct caataaaaaa tgtctggggt    2820
tgaaagaagt cgctcgcgtg gagagcttcc atggcttcat ctacggttgc ttcgaccagg    2880
aggcccctcc tcttatggac tatctgggtg acgctgcttg gtacctggaa cctatgttca    2940
agcattccgg cggtttagaa ctggtcggtc ctccaggcaa ggttgtgatc aaggccaact    3000
ggaaggcacc cgcggaaaac tttgtgggag atgcatacca cgtgggttgg acgcacgcgt    3060
cttcgcttcg ctcgggggag tctatcttct gctcgctcgc tggcaatgcg gcgctaccac    3120
ctgaaggcgc aggcttgcaa atgacctcca aatacggcag cggcatgggt gtgttgtggg    3180
acggatattc aggtgtgcat agcgcagact tggttccgga attgatggca ttcggaggcg    3240
caaagcagga aaggctgaac aaagaaattg gcgatgttcg cgctcggatt tatcgcagcc    3300
acctcaactg caccgttttc ccgaacaaca gcatgctgac ctgctcgggt gttttcaaag    3360
tatggaaccc gatcgacgca aacaccaccg aggtctggac ctacgccatt gtcgaaaaag    3420
acatgcctga ggatctcaag cgccgcttgg ccgactctgt tcagcgaacg gtcgggcctg    3480
ctggcttctg ggaaagcgac gacaatgaca atatggaaac agcttcgcaa aacggcaaga    3540
aatatcaatc aagagatagt gatctgcttt caaaccttgg tttcggtgag gacgtatacg    3600
gcgacgcggt ctatccaggc gtcgtcggca aatcggcgat cggcgagacc agttatcgtg    3660
gtttctaccg ggcttaccag gcacacgtca gcagctccaa ctgggctgag ttcgagcatg    3720
cctctagtac ttggcatact gaacttacga agactactga tcgctaacag acgagtcgac    3780
catgatgatc aatattcaag aagacaagct ggtttccgcc cacgacgccg aagagattct    3840
tcgtttcttc aattgccacg actctgcttt gcaacaagaa gccactacgc tgctgaccca    3900
ggaagcgcat ttgttggaca ttcaggctta ccgtgcttgg ttagagcact gcgtggggtc    3960
agaggtgcaa tatcaggtca tttcacgcga actgcgcgca gcttcagagc gtcgttataa    4020
gctcaatgaa gccatgaacg tttacaacga aaattttcag caactgaaag ttcgagttga    4080
gcatcaactg gatccgcaaa actggggcaa cagcccgaag ctgcgcttta ctcgctttat    4140
caccaacgtc caggccgcaa tggacgtaaa tgacaaagag ctacttcaca tccgctccaa    4200
cgtcattctg caccgggcac gacgtggcaa tcaggtcgat gtcttctacg ccgcccggga    4260
```

```
agataaatgg aaacgtggcg aaggtggagt acgaaaattg gtccagcgat tcgtcgatta   4320
cccagagcgc atacttcaga cgcacaatct gatggtcttt ctgtgattca gtgaccattt   4380
ttacaaatgg tcactgcaac cgcggtcacc attaatcaaa gggaatgtac gtgtatgggc   4440
aatcaacaag tcgtttcgat aaccggtgca ggctcaggaa tcggtctcga actggttcgg   4500
tcctttaagt cggccggtta ttacgtatcc gctctcgtac gaaacgagga gcaagaggcg   4560
cttctttgca aagagttcaa ggacgcactc gagattgtag tgggcgatgt ccgggaccac   4620
gcaacaaatg agaagctgat aaagcaaaca atcgatagat tcggtcatct tgattgtttt   4680
attgcaaatg ccggtatctg ggattacatg ctgagcatcg aagagccttg ggagaaaata   4740
tcgagcagtt ttgacgaaat attcgacatt aatgtcaaga gctatttcag tggcatcagt   4800
gccgccctgc cggaactgaa aaagactaac ggatcagtgg tgatgaccgc ttcggtgtcg   4860
tcccatgcgg tcggtggtgg tggttcttgc tacatcgcca gcaagcatgc ggtgctcggt   4920
atggttaagg ctttggccta cgaattggcc cccgaagttc gcgtgaacgc tgtttcgccg   4980
gggggcaccg tgacgtctct gtgcggttcc gcgagcgccg gtttcgacaa aatgcacatg   5040
aaagacatgc ccggcatcga cgatatgatc aaaggtctca cgcctcttgg gtttgcagcc   5100
aagcccgaag acgtggtggc accctatttg ttgctggctt cgcgaaagca aggaaaattc   5160
atcaccggca ccgtgattag cattgatggc ggtatggcgc tcggtcgcaa gtgagcttgt   5220
agccgatcag aagttataga cacatttcag gtgacgcccc atgaagacaa aactgtttat   5280
caataacgcc tggatcgatt ctagtgacca gcagaccttc gagcgcatac accccgtcag   5340
cagcgatgtg gtgactgaga gcgcaaacgc cacagtgacg gacgcgataa aggcggcgca   5400
agcggccgag gaggcgttca agacctggaa ggccgttgga ccttcagagc gtcgccgcct   5460
tctcctaaag gtcgccgatg tcatggaaag taaaacaccc aagttcatcg aagtgatggc   5520
catggaggtg ggagcttccg ccctttgggc cggattcaac gtccatgcgt ctgccaatgt   5580
gttccgagag gctgcctcgc tggctaccca aattcagggt gaaaccatcc caacggacaa   5640
agccgaaacg ctctcaatga cactacgtca gccggtcggc ccgatcctaa gcatcgttcc   5700
atggaacggc accgcagtgc ttgcggcacg agccatcgct tatccgctgg tctgtggcaa   5760
cactgtggtg ttcaaaggct ctgaatttag tcccgcgacg catgccctga tcacccagtg   5820
cgtgcaggaa gccgggctgc ccgctggcgt gctcaattac ctcaactctt cgcctgaccg   5880
ttcgcccgag atcgctgacg cactgatctc tgccaaggag atccgccgca tcaacttcac   5940
gggttccacc cgcgtgggca gcattatcgc gcagaaagcc gcgcaacacc tcaagcgctg   6000
cctgctggag ctcggcggca agtccccgct tattgttctg gatgatgcag acatcgatgc   6060
ggcggtcaag gcagcggtgt tcggtagctt cctgttccaa ggtcagatct gcatgtccac   6120
tgagcgcttg atcgttgatg agaagatagc cgacgaattt gtcgcaaaat ttgtcgaaaa   6180
aactaagcgc ttgagcgcag gcgacccgtg cgtaactggc gactgcatca tcggcccgat   6240
ggtctcgcca aattcgggtg agcggatcaa tggtttgttc aaagacgcga tcgacaaagg   6300
ggcaaaagtt gtttgcggcg gcttggccca aggtgcgctc atgccggcca cgatcctgga   6360
tcacgtcaaa tctgacatgc ggatttacga tgaggagacc tttggtccca tcaccgtggt   6420
aatccgttgt aaaggcgaag cagaggccgt ccgcattgcc aacgacagcg tctatggcct   6480
gtcgtcgggc gtatttggcc gcgacatcaa ccgcgctcta cgcgtgggta tgtccatcga   6540
atatggttct gtacacatca acggttcgac cgtccagaac gaggcgcagg ctccttacgg   6600
aggcaccaag aacaccggct acgggcgctt cgacggccgt gctgtaatcg acgagttcac   6660
```

```
agagatcaag tggctgacca tcgaaccttt cgagcagcaa tatcccttct gataagcact    6720
aactcccagg aatcaaacta tgagtaagca agctgcagtt atcgagctcg gatacatggg    6780
tatctcggtc aaggaccctg atgcgtggaa atcatttgcc acggatatgc taggtctgca    6840
agttcttgat gagggtgaga aggaccgttt ctatctgcgg atggattact ggcatcatcg    6900
gatcgtagtc catcacaacg gacaggacga cttggagtac ctaggctggc gtgtagccgg    6960
caagccggag ttcgaagctc tgggtcaaaa gcttattgat gccggttaca agatccgcat    7020
ctgcgacaaa gttgaggctc aggagcgtat ggtgttgggt ctgatgaaga cagaagatcc    7080
gggcggcaac ccgaccgaga tattctgggg cccccggatc gacatgagca acccgttcca    7140
tcccggtcgc cccctgcacg gaaagtttgt gaccggtgac caaggcttgg gccattgcat    7200
cgttcgccaa accgacgtcg cagaagctca taagttttat agcctgctgg gcttccgtgg    7260
ggacgtcgaa taccggattc cgttgcccaa cggcatgact gccgaactgt cgttcatgca    7320
ttgcaacgcc cgtgatcact ccattgcttt tggtgccatg cccgctgcca aacgactcaa    7380
tcacttgatg cttgagtaca cccatatgga agacttggga tacacgcacc aacagtttgt    7440
aaagaacgaa attgacattg ccttgcagct tggcattcac gccaacgaca aggcgttgac    7500
gttctatggt gcaacgcctt cgggctggct cattgagccc ggctggcgag gtgccacggc    7560
catagatgaa gcggagtatt acgtcggcga catcttcggc catggcgtgg aggccactgg    7620
atatggcctg gatgtaaaac tgagctaaag atgcgcgctc gttgggcgag gctctagtcc    7680
agcatcttca tacgcaacca accttgcagg gcgatgagat caaaggacgt taaagcgaag    7740
gggaagtggt tcgggccatg cgcataccga tccatgacat ttgtttcata gtatataggt    7800
agataggtga atcaagcgct tagtcaacta gtggacacat ctgttccatg aggctatcta    7860
ctatctattc aaaacaagaa taataaatag gatgaaaata ataatgataa aaagaacgat    7920
ttgtcttgtg tatcctctat tctgtttggc aagccccaca tgggccgaag agtcgccttg    7980
gacgtaccgt attggtatga ctaatgtagc tttcgatgct agcgcaaaag tatacttaaa    8040
tggtcagcgg gtgccaggag gaagcgctga tgcgagcgat aacaacgcgc ttacattcga    8100
cttcggctac gccatcaacg accagtggaa tgtacgtgcg attgtcggta ttccgcctac    8160
aactaaagtg acgggcgcag gcacacttcc tggtatccag ctggggaaaa taacttacgc    8220
tccaacagta ttaacgttga actataacct ccccgctttg ggtcccgttc gccctcacat    8280
aggtgcggga gtcaattaca cgcggatttt tgaaagtcgg gacgctaatc taaaatcgtt    8340
cgatgccgac cacgcttggt cccccgcgct acatgttggt gccgatattg acgttaaccg    8400
tggttggttc gttagcattg atatccggaa gttatacctg aaaagcgacg catcagggta    8460
cttggggcca caggaggcta aagcacgggt aactcttgac ccattactaa cttcgatcgc    8520
gatcggacgc caattctgat gattctgttt aaagttcttt atctatctaa ccgcaaaggg    8580
tgtttccatg tcgaataaaa ttatgaaaac gtcgcgtctt accgccgaag atatcaacgg    8640
cgcctggact ataatgccca caccctcgac gcctgatgct tctgattggc gcagcactgc    8700
caccgtggac ttagaagaga ctgcccgcat agttgaagag ctgattgcag ctggtgtcaa    8760
cggtattcta agtatgggta cttttggtga gtgcgccacg ttgacctggg atgaaaaacg    8820
tgattatgtc tcgacgattg tcgagaccat tcgtggtcgc gtgccttatt tctgtggcac    8880
gacagcctta aatacccgag aagtcatccg ccagacccga gagcttatcg atattggcgc    8940
caacggcact atgctcgggg tgccgatgtg ggtgaagatg gacctgccta cagcggttca    9000
gttctatcgt gatgttgcag atgcggtacc agaggctgcc attgcgattt acgccaaccc    9060
```

```
cgaagcattc aagttcgact tccctcgccc attctgggca gagatgtcca aaattccgca    9120
ggtagtgact gcgaagtatc taggcatcgg aatgcttgac ttggacctga gactggcacc    9180
caacatccgc ttccttcccc acgaagatga ctattacgcg gccgcacgca tcaatcccga    9240
gcgcataacc gcgttctggt caagcggggc catgtgcggc ccggctaccg ccatcatgtt    9300
gcgtgacgaa gtggtgcggg ccaagagcac cggtgactgg gccaaggcca aagccatctc    9360
cgatgatatg cgtgcagccg actcgacatt gtttccgcgt ggcgactttt cggagttctc    9420
gaagtataat atcgggcttg aaaaggcacg gatggatgcg gctggttggc tcaaggctgg    9480
gccctgccgt ccgccctaca accttgttcc agaagactac ctcgctggtg cacagaaatc    9540
aggcaaggct tgggccgcgc tgcacgctaa atacagtaat gaattgaagt agttcacctc    9600
cgcagacctg agtgacaggg tggcgcagac gctgagggtg caggaattaa gtgagctaaa    9660
gcacatttct tgcgccaggc attgccagat cagcaaagtt tgctga                   9706
```

```
<210> 6
<211> 2294
<212> DNA
<213> Artificial Sequence

<220>
<223> A modified DNA sequence

<221> unsure
<222> (186)...(186)
<223> n = a or t or g or c

<400> 6
agggcagckt ctgtatttgc ggcagcgaaa tgctccctaa attcctcatt taccccatct      60
gaggattgct ttatgacagt aaagtggatt gaagcagtcg ctctttctga catccttgaa     120
ggtgacgtcc tcggcgtgac tgtcgagggc aaggagctgg cgctgtatga agttgaaggc     180
gaaatntacg ctaccgacaa cctgtgcacg catggttccg cccgcatgag tgatggttat     240
ctcgagggta gagaaatcga atgccccttg catcaaggtc ggtttgacgt ttgcacaggc     300
aaagccctgt gcgcaccccgt gacacagaac atcaaaacat atccagtcaa gattgagaac    360
ctgcgcgtaa tgattgattt gagctaagaa ttttaacagg aggcaccccg ggccctagag     420
cgtaatcacc cccattccat ctttttttagg tgaaaacatg aattacaata ataaaatctt    480
ggtaagtgaa tttggtctga gccaaaagca cctgattcat ggcgatgaag aacttttcca     540
acatgaactg aaaaccattt ttgcgcggaa ctggcttttt ctcactcatg atagcctgat     600
tcctgccccc ggcgactatg ttaccgcaaa aatggggatt gacgaggtca tcgtctcccg     660
gcagaacgac ggttcgattc gtgcttttct gaacgtttgc cggcatcgtg gcaagacgct     720
ggtgagcgtg gaagccggca atgccaaagg ttttgtttgc agctatcacg gctggggctt     780
```

09843250 .042901

cggctccaac ggtgaactgc agagcgttcc atttgaaaaa gatctgtacg gcgagtcgct 840
caataaaaaa tgtctggggt tgaaagaagt cgctcgcgtg gagagcttcc atggcttcat 900
ctacggttgc ttcgaccagg aggcccctcc tcttatggac tatctgggtg acgctgcttg 960
gtacctggaa cctatgttca agcattccgg cggtttagaa ctggtcggtc ctccaggcaa 1020
ggttgtgatc aaggccaact ggaaggcacc cgcggaaaac tttgtgggag atgcatacca 1080
cgtgggttgg acgcacgcgt cttcgcttcg ctcgggggag tctatcttct cgtcgctcgc 1140
tggcaatgcg gcgctaccac ctgaaggcgc aggcttgcaa atgacctcca aatacggcag 1200
cggcatgggt gtgttgtggg acggatattc aggtgtgcat agcgcagact tggttccgga 1260
attgatggca ttcggaggcg caaagcagga aaggctgaac aaagaaattg gcgatgttcg 1320
cgctcggatt tatcgcagcc acctcaactg caccgttttc ccgaacaaca gcatgctgac 1380
ctgctccggt gttttcaaag tatggaaccc gatcgacgca aacaccaccg aggtctggac 1440
ctacgccatt gtcgaaaaag acatgcctga ggatctcaag cgccgcttgg ccgactctgt 1500
tcagcgaacg gtcgggcctg ctggcttctg ggaaagcgac gacaatgaca atatggaaac 1560
agcttcgcaa aacggcaaga aatatcaatc aagagatagt gatctgcttt caaaccttgg 1620
tttcggtgag gacgtatacg gcgacgcggt ctatccaggc gtcgtcggca aatcggcgat 1680
cggcgagacc agttatcgtg gtttctaccg ggcttaccga gcacacgtca gcagctccaa 1740
ctgggctgag ttcgagcatg cctctagtac ttggcatact gaacttacga agactactga 1800
tcgctaacag acgagtcgac catgatgatc aatattcaag aagacaagct ggtttccgcc 1860
cacgacgccg aagagattct tcgtttcttc aattgccacg actctgcttt gcaacaagaa 1920
gccactacgc tgctgaccca ggaagcgcat ttgttggaca ttcaggctta ccgtgcttgg 1980
ttagagcact gcgtggggtc agaggtgcaa tatcaggtca tttcacgcga actgcgcgca 2040
gcttcagagc gtcgttataa gctcaatgaa gccatgaacg tttacaacga aaatttttcag 2100
caactgaaag ttcgagttga gcatcaactg gatccgcaaa actggggcaa cagcccgaag 2160
ctgcgcttta ctcgctttat caccaacgtc caggccgcaa tggacgtaaa tgacaaagag 2220
ctacttcaca tccgctccaa cgtcattctg caccgggcac gacgtggcaa tcaggtcgat 2280
gtcttctacg ccgc 2294

<210> 7
<211> 4355
<212> DNA
<213> Artificial Sequence

<220>
<223> A modified DNA sequence

<400> 7
atcgataaaa ccatgagggt aataaataat gatgaaggag actatctaga attcgttctg 60
gagacggtgc cggtggcata accgtagacc agcaggctga gcaaaatacc tcggggtgat 120
aagccgcgct accgcgtccg ccatatgcct gagtgagtgc gctaaggtcg ggccgctcaa 180
tggcttcgag aataaagcgg gccaggtgat cttctgtcaa tctcgccgaa tgggtggaag 240

```
cacatcaacc tgaccggcga ttacgtctgg cggcagagcc gcagactgga ggtcgggaag      300
ttccggccgt tacggtggcc ctgaaaacct tagtgtacga ttttttccgt cttctctatt      360
cggcctgcgt tgcgacgttc gatctgacta gggcgcagcg gcaacatcgg ctgagtacga      420
tccagcgcct ggatctgggt tatttcgtca accgacaaca ccagagcgtt atccggtggg      480
ttcaggtaaa gcccaaccac attgactacc tctcggcaaa atgagggtcg ttactaattt      540
tgaacgtctt caagcgatgg ggttttaggt cggaagctgc ccagacctgc ccaacctacc      600
acgtcgtgac ccgagcatat tcggcccgcg gcggcaaaaa acacagaaat gagcgggtga      660
cccgatcgcc tttgatcgat tctccgcttt caaaatgggc gggggctgaa gtcagccaga      720
aatacatgac tacttcagac gattcgcagt attcacgctg gtgataaaca aattcaacta      780
tgctttattg acaaataaaa gcacgctcac catcatcgcg aatacaaatc ttataaaaat      840
taagccggat ttatggaact tctcatacag ccaaacaatc gcctcattag ctttagtccc      900
ggcgccaacc ttctggaagt gcttcgcgaa aacggtgtcg ctatttccta cagttgtatg      960
tctgggcgtt gcggaacctg ccgctgtcgg gttacagatg gcagtgtaat tgattcgggg     1020
gcgggaagcg ggttaccaaa cctcgtggac gagcattatg tgctcgcctg tcagtcagta     1080
cttactcaca attgcgcgat cgaaatccca gaaaccgacg aaatcgtcac ccacccggcg     1140
agaatcatca agggcactgt ggtcgccgtc gagtcgccca ctcacgatat ccgtcgccta     1200
cgcgtacgcc tcgctaagcc cttcgagttc tcacccggac agtacgcgac attgcagttc     1260
agtcctgagc atgcgcgtcc gtattcaatg gcaggtctgc cagatgacca agaaatggag     1320
ttccacatac gcaaggtgcc gggtgggcgc gtaacggagt atgttttcga gcacgtccgc     1380
gaaggtacaa gcatcaagtt gagcgggcca cttggtacgg cttatttgcg tcagaaccac     1440
accgggccga tgctctgtgt gggcggtggg accggactag caccggtgct gtcgattgtt     1500
cgcggcgcgc tgaagttggg tatgacaaac cccatcctcc tttatttcgg agtgcgcagt     1560
cagcaagacc tctacgacgc agagcgattg cacaaactcg ccgctgatca ccctcaactg     1620
accgtacaca cggtaatcgc aatgggcccg attaatgaga gtcagcgagc cggtctagtt     1680
accgatgtga tcgaaaaaga catcatttcg ctggctgggt ggagggccta cctgtgcggc     1740
gcaccagcga tggttgaagc gctttgcacc gttaccaagc atcttggaat atcacccgaa     1800
catatttatg ccgatgcctt ctatcccggt ggaatctgaa tagtcccttt ccctgcacct     1860
ctgtccatcg aggactcatc aggaggatac tcaaataagc gcaaataata acagtcgcgt     1920
cagtatttgc ggcagcgaaa tggtttccct ctccctcatt taccccatct gaggatagtt     1980
ttatgacaga aaaatggatt gaagcagtcg cccttttctga catcccagaa ggtgatgtcc     2040
tcggcgtgac tgtcgagggc aaggagctgg cgttgtacga agtggaaggc gaaatctacg     2100
ctaccgacaa cctgtgcacg catggtgccg cccgcatgag cgatggttat ctcgaggggc     2160
gagaaatcga atgccccttg catcaaggtc ggtttgacgt ttgtacaggc agagccctct     2220
gcgcccccgt gacagagaac atcaaaacat atgcagtcaa gattgagaac ctgcgcgtaa     2280
tgattgattt gagcggagag ttttaatagg aggtaccccg gaccctagcg cgtaactacc     2340
ccaattccat ctttttttagg tgaaaacatg aattacaaaa acaaaatatt ggtgagtgaa     2400
tctgggctga cccaaaagca cctgattcat ggcgatgaag aacttttcca gcacgaactg     2460
agaaccattt ttgcgcggaa ctggcttttt ctcactcatg acagcctgat tccatccccc     2520
ggcgactatg ttaccgcaaa aatgggtatt gacgaggtca tcgtctctcg gcagagcgac     2580
ggttcgattc gtgccttcct gaacgtttgt cggcaccgtg gcaagacgct ggttaacgcg     2640
```

```
gaagccggca atgccaaagg tttcgtttgc agctatcacg gctggggctt cggctccaac    2700
ggtgaactgc agagcgttcc attcgaaaaa gagctgtacg gcgagtcgct caacaaaaaa    2760
tgtctggggt tgaaagaagt cgctcgcgtg gagagcttcc atggcttcat ctatggttgc    2820
ttcgatcagg aggcccctcc tcttatggac tatctgggtg acgctgcttg gtacctagag    2880
cccatcttca aacattcagg cggtttagaa ctggtcggtc ctccaggcaa ggttgtgatc    2940
aaggccaact ggaaggcacc cgcggaaaac tttgtggggg atgcatacca cgtgggttgg    3000
acgcacgcgt cttcgcttcg ctcgggagag tctatcttcg cgtcgctcgc tggcaacgca    3060
gtgctgcccc ctgaaggtgc aggcttgcaa atgacctcca aatacggcag cggcatgggt    3120
gtgttgtggg acggatattc aggcgtgcat agcgcagact tggttccgga gttgatggca    3180
ttcggcggct ctaagcagga aaggctgaac aaagaaattg gcgatgttcg cgcccggatt    3240
tatcgcagcc acctcaactg caccgttttc ccgaacaaca gcatgctgac ctgctcgggt    3300
gttttcaaag tatggaaccc gatcgacgca aacaccaccg aggtctggac ctacgccatt    3360
gtcgaaaaag acatgcccga ggatctcaag cgccgcttgg ccgactcggt tcagcgtacg    3420
gttgggcctg ctggcttctg ggaaagcgac gacaatgaca atatggaaac agcgtcgcaa    3480
aacggcaaga aatatcaatc cagagatagt gatctgcttt caaaccttgg tttcggtaag    3540
gacgtatacg gcgacgcggt ctatcctggc gtcgtcggca aatcggcgat cggcgagacc    3600
agttatcgtg gtttctaccg ggcttaccag gcacacgtca gcagctccaa ctgggctgag    3660
ttcgaggatg cctctagtac ttggcatacc gaactgacga agactactga tcgctaacag    3720
acgagtcgac catgatgatc aatattcaag aagacaagct ggtctccgcc cacgacgccc    3780
aagagtttct tcgtttcttc aattgccacg acgcggcttt gcaacaagaa gccaccacgc    3840
tgctgaaccg ggaagcgcat ctgttggaca ttcaggctta ccgggcttgg ttagagcact    3900
gcgtggggtc agaggttcaa tatcaggtca tttcacgcga actgcgcgcc gcttccgagc    3960
gccgttataa gctcaatgaa gccatgaacg tttacaacga aaattttcag caactgaaag    4020
ttcgaatcga gcatcaactg gatccgcaaa actggagcaa cagcccgaag ctgcgcttta    4080
ctcgcttcat caccaatgtc caggccgcaa gggacgtaga tgacgaagag ctacttcaca    4140
tccgctccaa cgtcattctg caccgggcac gacgtggcaa tcaggtcgat gtcttctacg    4200
ccgcccggga agacaaatgg aaacgtggcg aaggtggagt gcgaaaattg gtccagcgat    4260
tcgtggatta cccagagcgc atacttcaga cgcacaatct gatggtcttt ctgtgatcca    4320
gtgaccactt ttacaaatgg tgactgctac cgcgg                               4355

<210> 8
<211> 2176
<212> DNA
<213> Artificial Sequence

<220>
<223> A modified DNA sequence

<400> 8
gtgatgtcct cggcgtgact gtcgagggta aggatctggc actgtacgaa gtggaaggcg      60
```

```
aaatctacgc taccgacgac ctgtgcacgc atggtgccgc ccgcatgagc gatggttatc     120
tcgaggggcg agaaatcgaa tgccccttgc atcaaggtcg gtttgacgtt tgtacaggca     180
gagccctctg cgcccccgtg acagagaaca tcaaaacata tgcagtcaag attgagaacc     240
tgcgcgtaat gattgattta agctgagaat ttttaatagg aggcgccccg gaccatagag     300
cgtaattatc cccattccat ctttttttag gtgaaaacat gaattacaaa aacaaaatct     360
tggtaagtga gtctgggctg acccaaaagc acctgattca tggcgatgaa gaacttttcc     420
agcacgaact gagaaccatt tktgcgcgga actggctttt tctcactcat gacagcctga     480
ttccttcccc cggcgactat gttaccgcaa aaatggggat tgacgaggtc atcgtctctc     540
ggcaaagcga cggttcgatt cgtgccttcc tgaacgtttg ccggcaccgc ggcaagacac     600
tggtgaacgc ggaagccggc aatgctaaag gtttcgtttg cagttatcac ggctggggct     660
tcggctccaa cggcgaactg cagagcgttc cattcgaaaa agagctgtac ggcgagtcgc     720
tcaacaaaaa atgtctgggg ttgaaagaag tcgctcgcgt agagagcttc catgggttca     780
tctatggttg cttcgatcag gaggcccctt ctcttatgga ctatctcggt gacgctgctt     840
ggtacctgga acccatcttc aaacattcag gcggtttaga actggtaggt cctccaggta     900
aggttgtgat caaggccaac tggaaggcac ctgcggaaaa ctttgtgggt gatgcatacc     960
acgtcggttg gacgcacgcg tcttcgcttc gcacaggcga atctatcttc tcgtcgctcg    1020
ctggcaacgc agtgctgccg cctgaaggtg cgggcttgca aatgacctcc aaatacggca    1080
gcggtatggg tgtgttgtgg gacggatatt caggcgtgca tagcgcagac ttggttccgg    1140
aattgatggc attcggcggc gctaagcagg aaaggctgaa caaagaaatt ggcgatgttc    1200
ccgcccggat ttatcgcagc cacctcaact gcaccgtttt cccgaacaac agcgtgctga    1260
cctgctcggg tgttttcaaa gtatggaacc cgatcgacgc aaacaccacc gaggtctgga    1320
cctacgccat tgtcgaaaaa gacatgcccg aggatctcaa gcgccgcttg gccgacgcgg    1380
ttcagcgaac ggtcgggcct gctggcttct gggaaagcga cgacaatgac aatatggaaa    1440
cagcatcgca aaacggcaaa aaatatcaat ccagagatag tgatctgatt tccaaccttg    1500
gtttcgggaa ggatgtatac ggcgacgcgg tctatcctgg cgtcgtcgga aaatcggcga    1560
tcggcgagac cagttatcgt ggtttctacc gggcttacca ggcacacgtc agcagctcca    1620
actgggctga gttcgaggat gcctctagta cttggcatac cgaactgacg aagactactg    1680
atcgctaaca gacgagtcga ccatgatgat caatactcaa gaagacaagc tggtatccgc    1740
ccatgacgcc gaagagtttc ttcgtttctt caattgccac gactcggctt tgcaacaaga    1800
agccaccacg ctgctgaccc gggaagcgca tctgctggac attcaggctt accggacttg    1860
gttagagcac tgcgtggggt cagaggttca atatcaagtc atttcacgcg aactgcgcgc    1920
cgcttccgag cgccgttata agctcaatga agccatgaac gtttacaacg aaaatttcca    1980
gcaactgaaa gttcgagtcg agcatcaact ggattcacaa aactggagca acagcccgaa    2040
gctgcgcttt actcgcttca tcaccaatgt ccaggccgca atggacgtaa atgatgaaga    2100
tctgcttcac gtccgctcca acgtcgttct gcaccgggca cgacgtggca atcaagtcga    2160
tgtcttctac gccgcc                                                    2176

        <210> 9
        <211> 14462
        <212> DNA
```

<213> Artificial Sequence

<220>
<223> A modified DNA sequence

<400> 9
```
gtcgactccg gtggccaaat cgacgggagc acaagcgtta cagcatcggg gctatcggcc      60
tcaaccgctg agacgcatct acatacccaa aagcaatggc aagaagcgtc cgctgggaat     120
tccaacgatg cgggatcgtg cgatgcaggc gctgtggaaa ctggcacttg agccggttgc     180
cgaaacgcgt gtagatccca actcctatgg atttcggccg cagcgatcca ctgccgatgc     240
aatcgcacat tgcttcaatg cactggcgaa acgtggttcg gcgcactggg tactggaagg     300
tgacattcga ggctgtttcg acaatatcag tcacgattgg ctgctcacca acgtgcccat     360
ggataaggtg gttctgcgca aatggcttcg agcggggtat gtcgatcagg gagccttgtt     420
cgcaacggag gcaggaaccc cgcaaggggg aatcatctct ccggtacttg cgaattggac     480
gctggatggc ctggaagatg tcgtccatgc aagcgtggct tcgacagcgc gcaagcgtaa     540
gccattcaag atacacgtcg tacgatatgc cgatgacttc atcatcacag ggccacgca     600
agctgttctg caacatcagg ttcgtcctgc aattgaggcg ttcctgaaac agcgtggact     660
ggaactctct gatgaagaga ctcagataac gcatatctca caaggcttcg attttctggg     720
ccagaacgta cgcaagtacg ccggcaagct actcatcact ccggcttgca agagtgtgaa     780
ggcactgctg gataaggtgc gggaaatcgc gaacgcaaac aaggcggcga ctcaagctaa     840
cttgatcttg accctgaacc cggtaatccg aggatgggcc atgtatcatc gccatgttgt     900
cgccgccaac gtttcgcatg gatcgaccat cagatttggc aagtgttatg gcgctgggca     960
gttcgtcggc acgccatgaa aaatgcctat tgggtaaaac aacgctactt tcgtgtcgta    1020
ggtcaacggc actgggtttt cgccactcag gagaaagcgc gcggcttgag tcaacccgct    1080
tggctgtatg ccgccgccag tgtttcgatc gtgcggcata tcaaaatatg cagtgcagcg    1140
aacccgttcg atccggcatg gacgttttac cttgagcgtc gaagagcgca tcgtcaggtt    1200
gcccagtctt actcaggctg ctggaaggct tgagccgtat gaggggtgac tcccaagtac    1260
ggttcttagg ggagacaggg gtggtaacac ccctgtttta cccgactgaa agcgcatatc    1320
ggagccgacg cagaatcggg cctggtacat catgtccatg cacagccgcc aacgtggctg    1380
acgtgacgga agttgcccat ctgttgcacg gtggcgagaa cgttgtctgt gcttggggtt    1440
acaccggcgt agagaaacgg cctgagcacg acggtcggca agtgatctgg cagatctgcg    1500
cgacgcagca cctacaagca tttgagcaaa cgcagtgcgc cgtacaaggc caggcgcaag    1560
atcgagaagg ccaaagccca ggtacgtgcg aaggtcgaac atccgtttcg ggtgatcaag    1620
cgtcagttcg gttatgtgaa gacccgcttc cgtgtcctgg ccaagaatac ggcggaattg    1680
accacactgt tcgcgctgtc gaactctgtg gatggtgcgc cggcaattat tgcctgctgc    1740
gggagaggtg cgcccgtgag tgacagaaaa ccagggcttt gcctcggtta tccaataaca    1800
cagcgactga aaatcgggca tttcggcatc cctaagccgc ccatttccga ctgatgagca    1860
acttgttcgg agtttcccta aagcacaagc gcaagccatc agaaaagagt gcctgagtgt    1920
gctccaagat agcccgaaac taaggccgga catggaaaag accaagatca cccatgtgaa    1980
tgatggactt gtttttgtgg gccatcggat cattcgcaag cgtagccgat acggagatat    2040
```

```
gcgtgtgatc acgacgatcc gagacaaagc cagaagattc gcagcctcac tgacggcgct    2100
gctatcaggt aaccacagtg aaagtaagat cgacatggtt gaatcagtca accgaaagct    2160
taaaggctgg gcggcgttta tcaattcgtt gattcaaagc caaagtattc agctatatcg    2220
atcgtgtcgt gttttggaaa ctggcccatt ggctggctca aaaataccgc tcccaggttt    2280
gtcttgcctg ttcctggccc gccgatgaac accacattct gcgcggtatc tgtgaactcc    2340
aggttggacc gctccctatg ggcatcaaca atatctcgtg gccttggctg ggatacgcaa    2400
ctaaccaaaa tgcgcgactg gcgcatgcag cttccatatg ggagacgctg gaagcgtcgg    2460
aatagaccgt caatcagaag accttcttcg ctcccgtcga actgaaagcc tcaccgagtg    2520
cgaacgtcaa cgggaaaggc gcgcaaagac ttgccttgct cacctcgtca cgagcgtaca    2580
aacgctgtga gctggctgta cgggtgcgct gcttagcgtt ggtttttccg ttctgtgagc    2640
caaccccgtg atatcgtagt gcaggaccat gccgttaagg agggctgcct ggtcagcatt    2700
ttcagccagc gccggaagtt cctgcacttc ggccttgaac tggcgttcca atgcctggat    2760
atggccatag ggaggtacgc tttggcgtaa ggcgttgcat gcagcttggt gctgtcgagg    2820
ctgatggtgc ggagcttgag cagcttcatt tgccctgcca actccaccgc ctgaaggaag    2880
ttcccggcca agtcatcgag gaagcgctgc cggaagctgg ccggggtgac tcttgcggcg    2940
aggtggcgaa aagcgagcga gtcgtaactg gcgcgctcaa tcctggagct gggaactttg    3000
ccggtggcat aaccatagac catcgggctg agcacaatgc ctcgggggtg ataagccgct    3060
accgcgtccg ccatatgcct gagtgagggc gctaaggtcg ggccgctcaa tggcttcgag    3120
aataaagcgg gccaggtgat cttctgccat ctctcccgaa tgggtggaag cacatcaacc    3180
tgaccagcga ttacgtctgc gcagagccgc agactagagg tcgggaagtt ccggccgtta    3240
cggctgcccg gaaaacctca gtgtacgatt ttttcagtct cctctattcg acctgcgttg    3300
cgacgttcga tctgactagg gcgcagcggc aacatcggct gagtacgatc cagcgcctgg    3360
atttgggtta tttcgttaac cgacaacacc agagcgttat cctgcgggtt caggcagagc    3420
ccaaccacat gactacctct cggcaaaatg agggtcgtta ctaattttga acgtcttcaa    3480
gcgatggggt tttaggtcga aagctgccca ggcctgcgga acctaccaca tcgtaacccg    3540
agcatattcg gcccgcggcg gctaaaaaca cagaaatgag cggggtgacc cgatcgcctt    3600
tgatcgattc tccgctttca aaagcggcag ggctgaagt cagccggaaa tacccagacc    3660
aatcacaaca ttcatgctgg tgataaataa attcaactat gctttattga caaataaaag    3720
cacactcacc atcatcgcga atacaaatct tataaaatta agccggattt atgaaacttc    3780
tcatacagcc aaacaatcgc ctcattagct ttagtcccgg cgccaacctt ctggaagtgc    3840
ttcgcgagaa cggtgtcgct atttcctaca gttgtatgtc tgggcgttgc ggaacctgcc    3900
ggtgccgggt tacagatggt agtgtaattg attcggggac gggaagcggg ttaccacacc    3960
tcgtggacga gcattatgtg ctcgcctgtc ggtcagtact tactaacaat tgcgcgatcg    4020
aaatcccaga agccgacgaa atcgtcaccc acccggcgag aatcatcaag ggcactgtgg    4080
tcgctattga gtcgcccact cacgatatcc gtcgcctacg cgtacgcctc gccaagccct    4140
tcgagttctc acccggacag tacgcgacat tgcagttcag tcctgagcat gcgcgtccgt    4200
attcaatggc aggtctgcca gatgaccaag aaatggagtt ccacatacgc aaggtgccgg    4260
gtgggcgcgt aactgagtat gttttcgagc acgtccgcga aggtacaagc atcaaattga    4320
gcgggccact tggtacggct tatttacgtc agaaccacac cgggccgatg ctctgtgtgg    4380
gcggtggaac cggactagca ccggtgctgt cgattattcg cggcgcgctg aagttgggta    4440
```

```
tgacaaaccc catcctcctt tatttcggag tgcgcagtca gcaagacctc tacgacgcag   4500
agcgattgca taacctcgcg gctgatcacc ctcaactgac cgtacacacg gtaatcgcaa   4560
tgggcccgat taatgagagt cagcgagccg gtctagttac cgatgcgatc gaaaaagaca   4620
tcagttcgct ggctgggtgg agggcctatc tgtgcggtgc accagcgatg gttgaagcgc   4680
tttgcaccgt taccaaacat cttggaatat caccggaaca tatttatgcc gatgccttct   4740
atcccggtga aatctgaatc gtccctttcc ctcacctcgg tccattgagg actcatcagg   4800
aggatactca aataggcgca aataataaca gccgcgtcac tatgtgcggc agcgaaatgg   4860
tttccctctc cctcatttgc ccatcggag gatagtttta tgacagaaaa atggattgaa   4920
gcagtcgccc tttgtgacat tccagaaggt gatgtcctcg gcgtgactgt cgagggtaag   4980
gatctggcac tgtacgaagt ggaaggcgaa atctacgcta ccgacgacct gtgcacgcat   5040
ggtgccgccc gcatgagcga tggttatctc gaggggcgag aaatcgaatg ccccttgcat   5100
caaggtcggt ttgacgtttg tacaagcaga gccctctgcg cccccgtgac agagaacatc   5160
aaaacatatg cagtcaagat tgagaacctg cgcgtaatga ttgatttaag ctgagaattt   5220
ttaataggcg gcgccccgga ccatagagcg tgattatccc cattccatct ttttttaggt   5280
gaaaacatga attacaaaaa caaaatcttg gtaagtgagt ctgggctgac ccaaaagcac   5340
ctgattcatg gcggtgaagg gcttttccag cacgaactga gagccgtttt tgcgcggaac   5400
tggcttttc tcactcatga cagcctgatt ccttcccccg gcgactatgt taccgcaaaa   5460
atggggattg acgaggtcat cgtctctcgg caaagcgacg gttcgattcg tgccttcctg   5520
aacgtttgcc ggcaccgcgg caagacactg gtgaacgcgg aagccggcaa tgctaaaggt   5580
ttcgtttgca gttatcacgg ctggggcttc ggctccaacg gcgaactgca gagcgttcca   5640
ttcgaaaaag agctgtacgg cgagtcgctc aacaaaaaat gtctggggtt gaaagaagtc   5700
gctcgcgtag agagcttcca tgggttcatc tatgcctgca tcgatcagga ggccccttct   5760
cttatggact atctcggtga cgctgcttgg tacctggaac ccatcttcaa acattcaggc   5820
ggtttagaac tggtaggccc tccaggtaag gttgtgatca aggccaactg gaaggcacct   5880
gcggaaaact ttgtgggtga tgcataccac gtcggttgga cgcacgcgtc ttcgctctgc   5940
acaggcgagt ctatcttctc gtcgctcgct ggcaacgcag tgctgccgcc tgaaggtgcg   6000
ggcttgcaaa tgacctccaa atacggcagc ggtatgggtg tgttgtggga cggatattca   6060
ggcgtgcata gcgcagactt ggttccggaa ttgatggcat tcggcggcgc taagcaggaa   6120
aggctgaaca aagaaattgg cgatgttcgc gcccggattt atcgcagcca cctcaactgc   6180
accgttttcc cgaacaacag cgtgctgacc tgctcgggtg ttttcaaagt atggaacccg   6240
atcgacgcaa acaccaccga ggtctggacc tacgccattg tcgaaaaaga catgcccgag   6300
gatctcaagc gccgcttggc cgacgcggtt cagcgaacgg tcgggcctgc tggcttctgg   6360
gaaagcgacg acaatgacaa tatggaaaca gcatcgcaaa acggcaaaaa atatcaatcc   6420
agagatagtg atctgatttc caaccttggt ttcgggaagg atgtatacgg cgacgcggtc   6480
tatcctggcg tcgtcggaaa atcggcgatc ggcgagacca gttatcgtgg tttctaccgg   6540
gcttaccagg cacacgtcag cagctccaac tgggctgagt tcgaggatgc ctctagtact   6600
tggcataccg aactgacgaa gactactgat cgctaacaga cgagtcgacc atgatgatca   6660
atactcaaga agacaagctg gtatccgccc atgacgccga agagtttctt cgtttcttca   6720
attgccacga ctcggctttg caacaagaag ccaccacgct gctgacccgg gaagcgcatc   6780
tgctggacat tcaggcttac cggacttggt tagagcactg cgtggggtca gaggttcaat   6840
```

```
atcaagtcat ttcacgcgaa ctgcgcgccg cttccgagcg acgttataag ctcaatgaag    6900
ccatgaacgt ttgcgacgaa aatttccagc aactgaaagt tcgagtcgag catcaactgg    6960
attcacaaaa ctggagcaac agcccgaagc tgcgctttac tcgcttcatc accaatgtcc    7020
aggccgcaat cgagctaaat gatgaagatc tgcttcacgt ccgctccaac gtcgttctgc    7080
accgggcacg acgtggcaat caagtcgatg tcttctacgc cgcccgggaa gacaaatgga    7140
aacgtggcga aggtggagtg cgaaaattgg tgcagcgatt cgtggattac ccagagcgca    7200
tacttcagac gcacaatctg atggtctttc tgtgatccgg tgaccacttt tacaaatggt    7260
gactgctacc gcggtcacca ttaatcaaaa gggaatgtac gtgtatgggc aatcaacaag    7320
tcgtttcaat aaccggtgcc ggctcaggaa tcggtctcga actggttcga tccttcaagt    7380
cggccggtta ttgcgtatcc gctctcgtac gaaacgagga gcaagaggcg cttctttgca    7440
atgaattcaa ggacgcactc gagatcgttg tgggcgatgt ccgagatcac gcaataaatg    7500
agaagctgat caagcagaca atcgctagat tcggtcatct cgattgtttc atcgcaaatg    7560
ccggtatttg ggattacatg ctgagcatcg aagagccttg ggagaaaatc tccagcagtt    7620
ttgacgaaat attcgacatc aatgtaaaga gctatttcag tggcatcagt gcagctctgc    7680
cggaactgaa aaagacgaac ggatcggtgg tgatgaccgc ttcggtgtcg tcccatgcgg    7740
tcggtggtgg tggttcttgc tacatcgcca gcaagcatgc ggtgttaggt atggtcaagg    7800
ctttggccta cgaattggct cccgaaattc gcgtgaacgc tgtctcgccg ggcggcaccg    7860
tgacgtctct gtgcggtcct gcaagcgccg gtttcgacaa aatgcacatg aaagacatgc    7920
ccggcatcga cgatatgatc aaaggcctca ctcctcttgg gtttgcagcc aagcccgaag    7980
acgtggtgga gccctatctg ttgctggctt cgcgaaagca gggaaaattc atcaccggca    8040
ccgtgattag cattgatggc ggtatggcgc tcggtcgcaa gtgagcttgc agccgatcaa    8100
aggttataga aacattttta ggtgacgccc catgaagaca aaactgttta tcaacaacgc    8160
ctggatcgat tccagtgacc agcagacctt cgagcgcaag caccccgtca acagcgaggt    8220
gatgactgag agcgcaaacg ccacggtgac ggacgcgata aaggcggcgc aagtggccga    8280
ggaggcattc aagacctgga aggacgttgg accttcggag cgtcgccgcc ttctcctgaa    8340
ggtcgccgat gtcatggaaa gtaaaacacc caagtttatc gaagtgatgg ccatggaggt    8400
gggagcttcc gctctttggg ccggattcaa cgtccatgcg tctgccaatg tgttccggga    8460
ggctgcctcg ctggccactc aaattcaggg cgaaaccatc ccaacggaca aagccgaaac    8520
gctctcaatg acactacgtc agccggtcgg cccgatcctg agcatcgtcc catggaacgg    8580
caccgcagtg ctggcggcac gagccatcgc gtatccgctg gtctgtggca acacggtggt    8640
gttcaaaggc tctgaattta gtcccgcgac gcatgccctg atcacccagt gcgtccagga    8700
agccgggctg cccgctggcg tgctcaacta tctcaactcc tcgcctgacc gttcgcccga    8760
gatcgccgac gcactgatct ccgcgaagga gatccgccgc atcaacttca cgggttcaac    8820
ccgcgtgggc agcattatcg cgcagaaggc cgcgcaacac ctcaagcgct gcctactgga    8880
gctcggcggc aagtccccgc ttattgttct ggatgacgca aacattgacg cggcggtcaa    8940
ggcagcggtg ttcggtagct tcctgttcca aggtcagatc tgcatgtcca ctgagcgctt    9000
ggtggttgat gagaagattg ccgacgaatt tgtcgccaag tttgtcgaaa aaactaagcg    9060
cttgagcgtg ggcgacccgt gcgtaactgg cgactgcatc atcggcccaa tggtctcgcc    9120
aaattcgggc gagcggatca atggtttgtt caaggacgcg atcgataaag ggccaaagt    9180
tgtttgcggc ggcatggccc agggtgcggt catgccggcc acgatcctgg atcacgtgaa    9240
```

```
atctgacatg cggatctacg atgaggagac ctttggtccc atcacagtgg tgatccgttg    9300
caaaggggaa gcagaggcca tccgcattgc caacgacagc gtttatggtc tgtcgtcggg    9360
cgtgtttggc cgcgacatca accgtgctct gcgagtgggt atgtcgatcg aatatggttg    9420
cgtacacatc aacggctcga ccgtccagaa cgaggcgcag gctccttacg gaggtaccaa    9480
gaacaccggc tacgggcgct tcgacggccg tgccgtggtc gacgagttca cagagctcaa    9540
gtggctgacc attgagccat tcgagcagca atatcccttc tgaaagcact aactccaagg    9600
aatcaaacga tgagtaagca agctgcagtt atcgagctcg gatacatggg catctcagtc    9660
aaggatcctg atgcgtggaa atcgtttgcc atgaatatgc tgggtctgca agtactcgat    9720
gagggtgaga aggaccgttt ctatctgcgg atggattact ggcaccatcg tatcgtagtt    9780
catcacggcg gagaggacga cttggaatat ctaggctggc gtgtagccgg caagccggag    9840
ttcgaagctt tggggcaaaa gctcattgat gccggttaca agatccgtgt ctgcgacaaa    9900
gttgaggctc aggagcgtat ggtgttgggc ctgatgaaga cagaagatcc aggcggcaac    9960
ccgaccgaga tattctgggg gccccggatc gacatgagca atccgttcca tcccggccgc   10020
cctctgcacg gcaagtttgt gaccggtgac cagggcttag gtcattgcat cgttcgccaa   10080
accgacgtcg ctgcggccca taaattctac agcctgctgg gcttccgtgg ggacgtcgaa   10140
taccgcattc cattgcccaa cggcatgact gccgaactgt cattcatgca ttgcaacgcc   10200
cgtgatcact ccattgcgtt tggtgccatg cctgctgcca agcggctcaa tcacctgatg   10260
cttgagtaca cccatatgga agacttggga tatacgcatc aacagttcgt gaagaacgaa   10320
attgacatcg ccttgcagct tggcattcac gccaacgaca aggcgttgac gttctacggc   10380
gcaacgcctt cgggctggct tatcgagccc ggctggcgag gtgctacggc catagacgaa   10440
gcggagtatt acgtcggcga catcttcggc catggcgtcg aggcacctgg atatggcctg   10500
gatgtaaaac tgagttaaaa gcgattatgc gtacatcggg tttctgacgt tttctgcttt   10560
ctatacagcg caataacaat aacaactaag tgggaatatt aataatgata aaaaaaacgt   10620
ttctttacat gcctcgcttt acttgttcac tcagcaccgc atgggccgaa gaatcacctt   10680
ggacataccg catcggcatg agtaatattg cttttgatac aagcgccaaa gtgtacttgg   10740
gaggccagcg tgtgccggga ggaagcgctg acgcgagcaa taacaacgcg ctcacattcg   10800
acttcggcta tgccatcaac gaccaatgga atgcacgttt gattgtcggt attccaccta   10860
caactaaagt taagggcgca ggcacacttc cgggcattca gctgggaaaa ataacttatg   10920
ctccaacatt actgacgtta aactataacc tcccagcttt tgggcccgtt cgtcctcaca   10980
tcggtgcagg agtcaattac acgcgaattt tggaaagcaa ggacgctaat ctaaaatcat   11040
tcgatgccga ccacgcttgg tcccccgcgc tgcatgttgg tgccgatatt gacgtgagtc   11100
gcaactggtt cgtcagcatt gatattcgga agttatacct gaaaaccgac gcatcaggtt   11160
acttagggcc acaggaggct aaagcaaagg taactcttga cccattgata acatcgatcg   11220
caatcggacg ccaattttga tgccccccttt taaggctctc tatctatcta actgcaaagg   11280
gtatttttat gttgaataaa attagtaaaa ccgcgcgtct taccgctgaa gatatcaatg   11340
gtgcctggac tataatgccc acaccgtcga cgcctgatgc ttctgattgg cgcacgacta   11400
acactgtgga cttagacgag actgcccgca tagttgaaga gctgattgct gctggtgtca   11460
acggtatttt gagtatgggt acctttggtg agtgcgccac gttgacctgg gaggagaaac   11520
gtgattatgt ttcgacggtt gtcgagacca ttcgcggtcg tgtaccttat ttctgcggca   11580
cgacggccct gaatacccga gaagtcattc ggcagacccg agagcttatc gatattggcg   11640
```

```
ctaacggcac catgctaggc gtgccgatgt gggttaagat ggacctgccc acagcggtcc   11700
agttctatcg tgatgttgca ggcgcggtac cggaggctgc cattgcgatt tacgccaacc   11760
ccgaagcatt caaattcgac ttccctcgcc cattttgggc agagatgtct aaaattcctc   11820
aggtagtgac tgccaagtat ctaggcatcg gaatgcttga cttggacctg aaattggcgc   11880
ctaacatccg cttccttcca cacgaggacg actattacgc ggccgcacgc atcaatcccg   11940
agcgcattac tgcgttctgg tcaagcgggg ccatgtgcgg cccggctacc gctatcgcgt   12000
tgcgtgatga agtggagcgg gccaagagta ccggtgactg gatcaaggcc aaagccatct   12060
ccgatgatat gcgtgcagcc gattcgacat tgtttccgcg tggcgacttt tcggagttct   12120
cgaagtataa tatcgggctt gaaaaggcac ggatggacgc ggctggttgg ctcaaggctg   12180
gtccctgccg tcctccctac aatcttgttc cagaagatta cctcgttggt gcacagaaat   12240
caggcaagtc gtgggccgcg ctgcacgcta aatacagtaa agaattaaag tagttcacct   12300
ccgcatgcct gagcgcgagg ggtgacgtaa acgccgagcg gtgcgggaag taagtgagtt   12360
agagttcatt tcttgtgcca ggcactgcta gatcagcaaa gttagctgat ctagcagtct   12420
cgaaaatttg ggcgaaagct gatcttagga atgcgggata aaggcagtac accgtaacga   12480
tcggggtgtg ccgttcatgt tgaacgacac cgctattgcg ccgacttctc ttcttcggag   12540
tgtttgattg tgattgtcga tttctatttc gatttttga gtccgttctc ttacttggcc   12600
aatcagcgtt tgtcaaagct tgcgcaagat catggcctta ccacgtgtta taacgcgatc   12660
gatttggcgc gggtcaaaat agcgatcggt aacgttggtc catctaatcg cgacttggaa   12720
gtaaaattgg actatttgaa agtagatttg caacggtggg cccagcttta cggaataccg   12780
ctggtatttc cagctaacta caacagcaga cggatgaata ctgggcttta ttactcggag   12840
gccgaggtgc aggccgctgc ctatgtgaat gtagtattta atgcgatttg ggagaaggc   12900
atagcaccag atttggaaag cttgcctgcc ctggtatctg aaaagctagg ctgggatcgt   12960
agtgccttcg agcgctttct cagcagcaac gccgcaacag agaggtatga cgagcagaca   13020
catttcgcca tcgagcgcaa ggtgttcggt gtgccaacga tgtttttggg cgatgaaatg   13080
tggtggggga acgaccgtct gtttatgctc gagagcgcaa tggggcgctt gtgccggaaa   13140
aatgccgatt taagtagttg atctgatcgt tatttgctcg atgagtcgct ttcaagatca   13200
gcggatactg aagtcagcta aatgcgggac tacttcaggc catgcttcag ggcggtcag   13260
ggcggactgg cgtgtgcacg gtagttgttg cgtgaggcat gtttttaaga tactattttc   13320
agtgcctgct gctctctctg ggttgagtgt gagggcgccg atagtctcaa atctcgcgtt   13380
gtacctgacc agagatacgg gtgtttatga tggcatcgct ggtacagcgc atgacgccat   13440
gctgcttcac ttcgagtgta ttcacgacct cgacgctagc aatctcaacg agccagatga   13500
tcgccagcca gggcggaagc tagtcgtgat tcgggttttc tacgttgtac ccgaatgcaa   13560
tcctggtctt aacggtccac gccttaaacg tggaatgcca aatctgcgat gtcgacattg   13620
gcccatataa tgttttggct ctcctctagt cagaaatcag taaggtgtgc atggtgataa   13680
ctggtcaagc cgggaatcag cggaatcgct gcgctggcgc cactggcgat caaacagcgg   13740
tcaaatacta ccattgaact ggcctaagag ggtgtggaca aataaagta aacttcttgc   13800
ctcctttcta tgcagcattg agccatgcct aaaactggac gcccaccgta atcgctacgc   13860
accattaccc tctgctgacc aggctcgctc atacgcagcc ttattccagt caagccgaac   13920
tggcgcaaga ttccatgcag aaacgggcat caccgcgcat cccgacacct ttgcaaaagc   13980
gctgagactg gccgggatcg ttcgagtaaa ggagcgtgcc aaaggtagct tccagcctcc   14040
```

```
cgagcctcgt aagtcttatg gctataccga ggcacaccgg cgtcagttgc cggagcaacg   14100
ctatcccagt tgcctgactg atgcagaatg gacgttggtc gctgatttat ttgaagtctc   14160
gggaggtcgc ggcgtaccgc tcgccactcc cgacgcactc ttttggaagc ctgctgttat   14220
gtcgtgcgca cagggtgctc ctggcgaatg ctgcttcgcg agtttcccta ttgggacaat   14280
gtctataaaa cctttcggcg ctggagcgtc agggcaagtt tgagcaaatg cacgaccggc   14340
tccgtgccca atggcgagag cgggtggatc gtgatgaacg gccgtcagcc gctgttctgg   14400
attcccaatc aacgcgcacg tctccccaag gcggtgaaag cggttacgac gcaggcaaaa   14460
ag                                                                  14462
```

```
<210> 10
<211> 12808
<212> DNA
<213> Artificial Sequence

<220>
<223> A modified DNA sequence

<400> 10
gtcgacgcat gatgcaggac tgggccgatc gcctcgacct cttcgagcag aaccaggtcg      60
aggcggccag catgccgctc accgtgcatc tggaaggtgt gcccgcattc ccgagtgagc     120
aaaccgcaag cgcgccctct acgccggttg ccgcttcgcc aagcctgctc gtgacgaagc     180
cgggtgacgc catgccgttg gtttctgccg ccgcacatcg gctgccggca gtgccgcccc     240
agcgatcggc cgcgccgctg gtgccttcgg acattcagcg cgagagaatg gatttgttcg     300
atgtcttcga agcgccgcac aaccttcccg tcgctgcgtt tgccaagatg gcgggcaaat     360
ctcgcaggtg gatcagctac gagatccagg cgggcaactt gctggcgttg aacgtgggca     420
accgcggcca gcgcgtgccg gactggcacc tcgacccgct caagcacgag ctgatccagt     480
ctgtcctgaa gctgaccagg ggtgcggacc cttggcagat ctaccatgca ctgctgcagc     540
cgcgctcgat gctgcggggg cgctcggcac tggagggcgt gactgccagc aatctcgaca     600
agctcgtcat ggccgtgagc acagcggtga aggaaagcga atggaccccg ctgcgggtcg     660
gtggtcgtgt agtcaaacaa cgggccggta agggatggcc tggacattat tcatattagt     720
gatactaata ttcatttatg gtttattgac tattagaata cagtcaatga tcatggaacc     780
ataaagcata taaataaaga agctagatct atggaacttc tcgtactacc gaacaatcgc     840
cgcttgcctt ttgattccgg tgccaacctt ttggaagtgc tccgtgagca ccgtgtgggt     900
atttcctaca gctgtatgtc tggacgatgc ggtacttgcc gctgccgagt tatagatggc     960
agcgtcatta gttcggcggc gaaaagcggt gactcaaatc gcatcgaaga gcattatgta    1020
ctcgcctgtc agtcagtgct caccagcaat tgcgcaattg agatcataga ctcagacgac    1080
atagtcactc acccggcgcg aatcatcaaa ggcatggttg tcgccgtcga gtcgcccact    1140
cacgatattc gccgcatccg cattcgcctc gccaagccct tcgagttctc acccggacag    1200
```

09845250 .042501

```
tacgcgatgc tacagttcag tcccgaacat gtgcgtccat attcaatggc tggtctgcca   1260
gatgaccaag aaatggagtt ccatatccgc aaagtgccgg gcgggcgtgt cacggagtat   1320
attttcgagc acgtccgcga aggtacaagc attaagttga gtgggcctct tggtacggcc   1380
tatctgcgtc aggctcacac cgggccgatg ctgtgtgtgg gcggcgggac cggactcgca   1440
ccggtgctgt cgattgttcg cggcgcgctg aagtcgggaa tgacgaaccc catccacctt   1500
tatttcgggg tgcgcagtca gcaagacctt tacgacgcag accgattgaa ccaactcgcg   1560
gctatccacc ctcaactgac tgtccataca gtgatcgcga cgggcccgat taatgagggc   1620
cagcgggccg gcctaattac cgatttgatt gaaaaagaca ttccctcgct ggctgggtgg   1680
agagcctacc tgtgcggcgc accagcgatg gttgacgctc tatgcaccgt cgccaaagat   1740
cttggaatat cgcccgagca tatttatgcc gacgccttct atcccagcgg ggtctgaatt   1800
gccccgaccc ttcacctctg tacatcgaga attcatcagg aagacactta aatgagcatc   1860
actaacaaca gccgcgtctg aatatttagg acagctggat gatctctaac tccatcatta   1920
ccccatttga agatagcttt atgacagaaa aatggattga cgcagtcgct ctttatgaaa   1980
tccctgaagg tgacgtcctc ggcgtgacag tcgaaggtaa ggaactagcg ctgtatgaag   2040
tggaaggcga aatctacgct accgacaacc tgtgcacaca tggtgctgcc cgcatgagtg   2100
atggctttct agaaggcaga gaaattgaat gtcctttgca tcaaggtaga tttgatgttt   2160
gcacaggcag ggccttgtgc gcccctgtga cacagaacat caaaacatac ccggtgaaga   2220
ttgagggcca gcgtgtgatg attgatttga gctgagaatt ttaataggag gcaccccgga   2280
ccctagagcg taatcccccc cattcgatct cttgaggtga aaatatgaat tacaaaaata   2340
aaaacttggt gagtgaatct gggctgaccc aaaaacacct gattcatggc gacgaagaac   2400
ttttccagcg cgaactggaa accatttttg ctcggaactg gcttttcctg actcatgaca   2460
gcctgattcc gtcccctggc gactatgtta cggcaaaaat gggggttgat gaggttatcg   2520
tctccaggca gaacgacggt tcgattcgtg cttttctgaa cgtttgtcgt caccgtggca   2580
agacgctggt acacgcagaa gcaggtaatg ctaaaggttt cgtttgcagc tatcacggct   2640
ggggcttcgg cgctaacggt gaactgcaga gcgtcccgtt tgaaaaagaa ctgtatggcg   2700
aggcgctcga caagaaatgt atgggattga aagaagtcgc tcgtgtagag agcttccatg.  2760
gcttcatcta tggttgcttc gatgaggaag ccccttctct caaagactac atgggggacg   2820
ctggctggta cctggagcct atgtttaagc attccggagg gctagaactg atcggtcctc   2880
caggaaaggt cataatcaag gctaactgga aagcgcccgc ggaaaacttt acgggggatg   2940
cgtaccacgt gggttggacg catgcgtctt cgcttcgctc agggcagtcg gtcttctcgt   3000
cgttagctgg caacgcagct ttgcccccag aaggtgcagg tctgcaaatg acctccaaat   3060
acggcagcgg catgggtgtg ttgtgggacg gatattcagg cgtgcacagc gcagacctgg   3120
ttccggaatt gatggccttc ggcggtgcta agcaggaacg gctgaacaaa gaaattggcg   3180
aggttcgcgc acgaatctat cgcagccacc tcaactgcac cgttttcccg aacaacagtt   3240
ttctgacctg ctcgggtgtc ttcaaggtat ggcacccgat cgacgcaaat accactgagg   3300
tatggaccta cgccatggtc gaaaaagaca tgcccgagga tctcaagcgc cgcttggtcg   3360
acgcggttca gagaacggtt gggcctgctg gcttctggga aagcgacgac aacgacaata   3420
tggaaacggt atcgcaaaac gccaagaaat atcagtccag agatggcgat ctggtttcca   3480
acctgggttt cggcggggac gtatacggcg acgaggttta tcctggcatc gtcggcaaat   3540
cggcgattgg cgagaccagt tatcgtggct tctatcgggc ttacggcgcg cacatcagca   3600
```

```
gctctagctg ggctgaattc gaggatgtct ctaaaaattg gcataccgaa ctggcaaaga    3660
ctactgatcg ctaacdgacg agagggacca tgatgattaa tattcaggaa gacaagcttg    3720
tctccgccca cgacgccgaa gagtttcttc gtttcttcaa ttccggcgac gaggctttgc    3780
aacaagaagc taccacgttg ctaacccggg aagcgcatct tttagacatt caggcttacc    3840
gcgcctggtt agagcactgc gtggactcag aggtgaaata tcagattatc tcacgcgaac    3900
tgcgctcagc ttccgagcgc cgttaccagc tcaatgaaac catgaacatt ttcaacgaga    3960
attatgaaca actggaagtt cgcgtagcgc atcaactgga tccgcaaaac tggggcaata    4020
gtccaaaggt gcgctttact cgtttcatca caaatatcca ggctgcaatg gacgaaaatg    4080
aagatttgct tcacattcgc tccaacctaa ttgttcaccg agcacgacgc ggcaatcaag    4140
tcgatgtctt ctatgccact cgggaggata aatggaagcg cggcgaagat ggagcgcgta    4200
agttggtcca acgattgatt gattatccag agcgcacatt ccagacgcac aatgtgatga    4260
tctttatgtg acccaataat cgcctttaca aatggtgact gctacaagcg gtcccattgt    4320
tcaaaaggaa atttatgtgt atgagcaatc aacaagtcgt ttcgataacc ggtgctggct    4380
caggaattgg tctcgaactg gttcgatcct ttaagtccgc cggttattgc gtatccgctc    4440
tcgtacaaaa cgaggagcaa aaggcgagcc tttgcaatga gttcaaggac gcactcgaga    4500
tcgtcgtggg cgatgtccgg gaccacgcaa caaatgagaa gctgataaag caaacaaccg    4560
atagattcgg ccatctcgat tgtttcattg caaatgccgg tatttgggat tacatgcttg    4620
gcatcgaaga gccttgggag aaaatatcga gcagttttga tgagatattc aacatcaatg    4680
tcaagagcta tttcagcggt atcagggccg ccctgcagga actgaaaaag actagcggat    4740
cagtggtgat gaccgcttca gtgtcgtccc atgcggtcgg tgctggtggt tcttgctaca    4800
tcgccagcaa gcatgcggtc ctgggcatga tgaaagcttt ggcttacgaa ttggctcccc    4860
acattcgcgt caacgccgta gcaccgggcg gcactgtgac gcctctgagc ggtcccgcaa    4920
gcgccggctt cgacaaaact cacatggaaa acatgcccgg tatcgaggac atgatcaagg    4980
gtctaacgcc tcttggaatt tcagccaagg ccgaagacgt agtggcaccc tatttgttgt    5040
tggcgtcgcg agatcaaggg aaattcatta ccgggactgt cattaatata gatggaggga    5100
tggcgctcgg tcgcaagtag gtttgtcgcc tatcttgaaa taataactaa atttctggta    5160
aaaccgcatg aatacaaaat tgtttatcaa caatgtctgg atcaattcca gtgaccaaca    5220
gaccttcgag cgaaagcacc ccgtcagtgg tgaggtgatg acggagtgtg caaactccac    5280
ggtgatggat gcgttaaagg ccgcgcaagc tgcccaagag gctttccaga cctggaagac    5340
tgttggacct tcggagcgtc gccgccttct gctgagggtc gctgaggtta tggaaagtaa    5400
aacacccgag tttatcgaag tgatggccaa ggaggtggga gcctccgctc tttgggccgg    5460
cttcaatgtc cagatgtcag ccaatgtgtt ccgtgaagcg gcatcgctgg ctacacaaat    5520
tcagggggaa actattccga cagacaagtc tgacacgctc tcaatgacgc tacgtcagcc    5580
ggtcggtccg atcctgagca tcgtgccgtg gaacggcacc gcagtgctgg cggcacgagc    5640
catcgcttat ccgctggtct gcggcaacgc ggtggtattc aaaggttctg agtttagtcc    5700
cgcgacgcat gccctgatca cccagtgcgt gcaggaagcc gggctgcctg ctggcgtgct    5760
caactatctc aactcttcgc ctgaccgttc gcccgagatc gccgacgcac tgatctcagc    5820
caaggagatc cgacgcatca acttcacggg ctccacccgc gtgggcagta ttatcgcgca    5880
gaaggccgcg caacacctca agcgctgcct gctggagctc ggtggcaagt ccccacttat    5940
tgttctggat gatgcagaca tcgatgcggc ggtcaaggca gcggtgttcg gtagcttcct    6000
```

09843250.042601

```
gttccaaggt cagatctgca tgtccactga gcgcttgatc gttgatgaga agatagccga   6060
cgaatttgtc gcaaaatttg tcgaaaaaac taagcgcttg agcgcaggcg acccgtgcgt   6120
aactggcgac tgcatcatcg gcccgatggt ctcgccaaat tcgggtgagc ggatcaatgg   6180
tttgttcaaa gacgcgatcg acaaaggggc aaaagttgtt tgcggcggct tggcccaagg   6240
tgcgctcatg ccggccacga tcctggatca cgtcaaatct gacatgcgga tttacgatga   6300
ggagaccttt ggtcccatca ccgtggtaat ccgttgtaaa ggcgaagcag aggccgtccg   6360
cattgccaac gacagcgtct atggcctgtc gtcgggcgta tttgggcgcg acatcaaccg   6420
cgctctacgc gtgggtatgt ccatcgaata tggttctgta cacatcaacg gttcgaccgt   6480
ccagaacgag gcgcaggctc cttacggagg caccaagaac accggctacg ggcgcttcga   6540
cggccgtgct gtaatcgacg agttcacaga gatcaagtgg ctgaccatcg aacctttcga   6600
gcagcaatat cccttctgat aagcactaac tcccaggaat caaactatga gtaagcaagc   6660
tgcagttatc gagctcggat acatgggtat ctcggtcaag gaccctgatg cgtggaaatc   6720
atttgccacg gatatgctag gtctgcaagt tcttgatgag ggtgagaagg accgtttcta   6780
tctgcggatg gattactggc atcatcggat cgtagtccat cacaacggac aggacgactt   6840
ggagtaccta ggctggcgtg tagccggcaa gccggagttc gaagctctgg gtcaaaagct   6900
tattgatgcc ggttacaaga tccgcatctg cgacaaagtt gaggctcagg agcgtatggt   6960
gttgggtctg atgaagacag aagatccggg cggcaacccg accgagatat tctggggccc   7020
ccggatcgac atgagcaacc cgttccatcc cggtcgcccc ctgcacggaa agtttgtgac   7080
cggtgaccaa ggcttgggcc attgcatcgt tcgccaaacc gacgtcgcag aagctcataa   7140
gttttatagc ctgctgggct tccgtgggga cgtcgaatac cggattccgt tgcccaacgg   7200
catgactgcc gaactgtcgt tcatgcattg caacgcccgt gatcactcca ttgcttttgg   7260
tgccatgccc gctgccaaac gactcaatca cttgatgctt gagtacaccc atatggaaga   7320
cttgggatac acgcaccaac agtttgtaaa gaacgaaatt gacattgcct tgcagcttgg   7380
cattcacgcc aacgacaagg cgttgacgtt ctatggtgca acgccttcgg gctggctcat   7440
tgagcccggc tggcgaggtg ccacggccat agatgaagcg gagtattacg tcggcgacat   7500
cttcggccat ggcgtggagg ccactggata tggcctggat gtaaaactga gctaaagatg   7560
cgcgctcgtt gggcgaggct ctagtccagc atcttcatac gcaaccaacc ttgcagggcg   7620
atgagatcaa aggacgttaa agcgaagggg aagtggttcg ggccatgcgc ataccgatcc   7680
atgacatttg tttcatagta tataggtaga taggtgaatc aagcgcttag tcaactagtg   7740
gacacatctg ttccatgagg ctatctacta tctattcaaa acaagaatga taaataggat   7800
gaaaataata atgataaaaa gaacgatttg tcttgtgtat cctctattct gtttggcaag   7860
ccccacatgg gccgaagagt cgccttggac gtaccgtatt ggtatgacta atgtagcttt   7920
cgatgctagc gcaaaagtat acttaaatgg tcagcgggtg ccaggaggaa gcgctgatgc   7980
gagcgataac aacgcgctta cattcgactt cggctacgcc atcaacgacc agtggaatgt   8040
acgtgcgatt gtcggtattc cgcctacaac taaagtgacg ggcgcaggca cacttcctgg   8100
tatccagctg gggaaaataa cttacgctcc aacagtatta acgttgaact ataacctccc   8160
cgctttgggt cccgttcgcc ctcacatagg tgcgggagtc aattacacgc ggatttttga   8220
aagtcgggac gctaatctaa aatcgttcga tgccgaccac gcttggtccc ccgcgctaca   8280
tgttggtgcc gatattgacg ttaaccgtgg ttggttcgtt agcattgata tccggaagtt   8340
atacctgaaa accgacgcat cagggtactt ggggccacag gaggctaaag cacgggtaac   8400
```

```
tcttgaccca ttactaactt cgatcgcgat cggacgccaa ttctgatgat tctgtttaaa    8460
gttctttatc tatctaaccg caaagggtgt ttccatgtcg aataaaatta tgaaaacgtc    8520
gcgtcttacc gccgaagata tcaacggcgc ctggactata atgcccacac cctcgacgcc    8580
tgatgcttct gattggcgca gcactgccac cgtggactta gaagagactg cccgcatagt    8640
tgaagagctg attgcagctg gtgtcaacgg tattctaagt atgggtactt ttggtgagtg    8700
cgccacgttg acctgggatg aaaaacgtga ttatgtctcg acgattgtcg agaccattcg    8760
tggtcgcgtg ccttatttct gtggcacgac agccttaaat acccgagaag tcatccgcca    8820
gacccgagag cttatcgata ttggcgccaa cggcactatg ctcggggtgc cgatgtgggt    8880
gaagatggac ctgcctacag cggttcagtt ctatcgtgat gttgcagatg cggtaccaga    8940
ggctgccatt gcgatttacg ccaaccccga agcattcaag ttcgacttcc ctcgcccatt    9000
ctgggcagag atgtccaaaa ttccgcaggt agtgactgcg aagtatctag gcatcggaat    9060
gcttgacttg gacctgagac tggcacccaa catccgcttc cttccccacg aagatgacta    9120
ttacgcggcc gcacgcatca atcccgagcg cataaccgcg ttctggtcaa gcggggccat    9180
gtgcggcccg gctaccgcca tcatgttgcg tgacgaagtg gtgcgggcca agagcaccgg    9240
tgactgggcc aaggccaaag ccatctccga tgatatgcgt gcagccgact cgacattgtt    9300
tccgcgtggc gacttttcgg agttctcgaa gtataatatc gggcttgaaa aggcacggat    9360
ggatgcggct ggttggctca aggctgggcc ctgccgtccg ccctacaacc ttgttccaga    9420
agactacctc gctggtgcac agaaatcagg caaggcttgg gccgcgctgc acgctaaata    9480
cagtaatgaa ttgaagtagt tcacctccgc agacctgagt gacagggtgg cgcagacgct    9540
gagggtgcag gaattaagtg agctaaagca catttcttgc gccaggcatt gccagatcag    9600
caaagtttgc tgatctggca gtttcaaaaa tttgggcgaa agctgatatc aggaatacgg    9660
gataaaggca gtgcaccata acgacggggg cgtgccattc gtgatgaacg attttgctat    9720
tgtgccgact tctgttcttg gagtgtttga ttgtgattgt cgatttttat ttcgattttt    9780
tgagtccgtt ctcttacttg gccaaccatc gtttgtcaaa gcttgcgcaa gactatggct    9840
tttccattcg ttattacgca atcgatttgg cgcgagttaa aatagccatc ggaaacgttg    9900
gtccatctaa tcgcgacctg atagtcaagc tggactattt gaaagtagat ttgcaacggt    9960
gggccgagct ttacgaaata ccgttggtat tcccagctaa ctacaacagc cgacggatga    10020
atactgggct ttattactcg ggagccatgg cacagactgg tgcctatgtg aatgtagtat    10080
ttaatgcggt ttggggagat ggcatagctc cagatttgga aagcttgcct gctctggtat    10140
ctgaaaaact aggctgggat cgtagcgcct tcgaggactt tatcagcagc gatgccgcaa    10200
cagagaggta tgacgagcag acacatgccg cgatcgaacg caaagtgttc ggtgtgccaa    10260
cgatgttttt gggcgatgaa atgtggtggg gaaacgaccg tctatttatg ctcgagaacg    10320
cagtgggagg tgcgcctgta aatggagaat agtcgctacg gagcgcttgt gccggctaaa    10380
tgccgatata agtggttgac ctgatcgtta tttgctcgat acagcgcttt caaaatcagc    10440
ggctactgaa gtcagataaa aatgcgggac tacttcaggc atcctgtgcg acacaaagtt    10500
ttacctgtaa ttgtccacct attccgagtt tggaatggta gctgactcgc tatgcgacca    10560
gcgatagcct aacaagacat gcatcactgg taacggtggg gtgtgaagct cctgcaacaa    10620
tgtagcccct tgatgtgtgt atttgctgcg aggtgaagca cagatgctcg gagccgtacc    10680
ggcttgtggc gctaggctgg caagtatgag caacgtaagt gggggttggg gcgcaatggg    10740
aaccaaaaac caacgcaagc cttaccagcg tcgttcggtg ccttcctccc atgcctccgc    10800
```

```
ctcgataaag cagctgcgca tatcggcttc ctggctgatc tcggttagta ggtcatgcaa   10860
ggtcttgtcc agcgcctcgt cgctccgata cggaatggtc agctcgtaat ggccggtctc   10920
cgaccgcttc atgccgtagg gctccaggca gtagcgctcg atgttctccg tggcccgctt   10980
ccgaccgcgc atgaacttgc tgttgttcac caccgccagg cgcagggtga cggtggccac   11040
ccgctcgacg gttgactctg ccggtgacgc gatattgcgc ttttgacctc gcgccagggc   11100
gctcttctgg tacgtcccga tctcgacgcc acggtggcgt aggtagctgt acagggtgct   11160
cttggagatg tgcaacttct cgccgatggc gctgacgctc aggcggccct cgcggtacag   11220
ggtctccgcc gccatggcgg tggcctcggc cttggctggc aggcccttgg gacggcgacc   11280
gatccggcct cgagtccgtg ccgccgacag gcccgcctga gtccgctcgc ggatcagctc   11340
gcgctcgaac tgaacaggtt gaacaccagg cgatcttggg cgtgggtgct gtcaatgggg   11400
tcgttcaggc tctgcaagcc gactttgcgt gcagccagct agccgaccaa ctcaaccagg   11460
tgcttgagcg agcgaccgag gcgatccagc ttccagatca ccacggcatc gcccgctcga   11520
acatgggcta gcaacttgtc caactccggc cgcgcgcttt ttgcgccgct ggcgatgtct   11580
tgatagatgc gttcgcaccc ggcctgtttc agggcatcga cctggaggtc ggcgttctaa   11640
tcccgagtgc tcacccgcgt ataaccgatc ttcataaaaa gtaccgttta cttgactgca   11700
ttagtaatag ttgaactttg attaagctta ccagttattt gaaccctagc gcaggtgtaa   11760
gcgtccagcc gccccacctc tactcagctt gatgaaccga ggggcagtag ttcatcaatc   11820
cggctgttag ggaaactctg aagaagactt cctgattttg gcaaaatgcc cggatttcac   11880
ccgccgagtt ttccaatgaa gcagatgacc ttcctcgacg ccgagtatgc cggtaagcgc   11940
aaacagaccc gcaaagagct gttcctgatc gagatggatc gggtggtgcc gtggaagggt   12000
ttgattgcct tgatcgatcc gcattacccc aagggtgagg gtggcccgcc agcctatcca   12060
ctgacggcaa tgctgcgggt tcatctgatg cataagtggt tcggctacag cgacccggcg   12120
atggaggaag cgctgtacga gacgaccatc ctgcgccagt ttgccgggtt gagcctggag   12180
cgcattccgg actaaaccac catccccaac ttccgccggc tgctggagaa acacgaacta   12240
cctaccggca tcatgtttgc tgaattttcg cctttggctt tgagtgcggc attcaaccga   12300
tacaggtgtg gcatgacacg ccaaagaaat caaccttggc ggctagaaca cttgtcgatg   12360
aagcgaacaa tccgcaaacg ttaaggctac caggaaaaac ccaggatgcg gtagccgata   12420
gcccgcgcgc gcattggtct ccttccgtcc aagagtgtca aaggatattg gactaatgca   12480
gcatgcctcg accacatggc cgtggtccgt tcatctgttg gccatacgtc tatcgaatcg   12540
gacgcgggcg gcgacgacct cgggttgttg gcgttcggcc cagtcgatga gctcagtcag   12600
ggacggcatc agcgaagtcc cgagcggggt gagcccgtag cgcacagatg gggtgtcga    12660
gggcgttacc tcgcgccaga tgagtccgtc gcgttcgagg tgacgcagcg tctcggtcag   12720
catccggcgc gaaatgtcgg gcacggcgcg ggccagcgca ttgaatcgtt gcggaccctg   12780
cgacaaggtg accaggatca gcgtcgac                                      12808
```

<210> 11

<211> 15024

<212> DNA

<213> Artificial Sequence

<220>
<223> A modified DNA sequence

<400> 11

```
atcgtcctta agcgccgcac gcagccgttt aagtgaattg ctgacggcag gctgcgtcag      60
ccccagtttt tcgccggccg tcgatacgct ccggtcgagc agtagctggt tgaagaccac     120
cagcagattc aagtcgatgt cgcgcagatc catgatgcct caccattatt catgctggtg     180
attttaacta tcagacttga tctatagcgc tataccgatc gacgcgccag aatcgcagcc     240
attcggagac aactgaaaaa agagcttgca tggaactggt agtagaaccc ctcaatttgc     300
atctgaacgc ggagaccggc agcaccctgc ttgacgtgct caggtccaac gaggtcccca     360
tttcttatag ctgcatgtcg ggccgctgcg gcacttgccg ttgccgtgtg attgccggcc     420
atcttcgcga taacggcccc gagacagggc gcccgcaggc aggaaaggga acctacgttc     480
tggcctgtca ggcggttctg accgaagact gcacgatcga gattcctgaa tctgacgaga     540
tcgtggttca cccggcgcgc atcgtcaagg ggacggtcac agcgatagac gaagccaccc     600
atgacatccg gcgcctgcgc atcaaactgg ccaaaccgct tgagttcagc cctggccagt     660
acgcaacggt gcagttcacg cccgaatgcg tccgcccata ttcgatggcc gggctgccta     720
gcgatgcgga aatggagttt cagattcgcg cggttccggg cgggcatgtc agcaactacg     780
ttttcaatga actgtccgta ggcgcttcgg tgcggatcag cggccccctc ggaacggctt     840
atctgcggcg cacgcacacc ggccccatgc tttgtgtggg gggtggaaca ggtctggcgc     900
ccgtcctttc gatcgttcga ggcgcactgg aaagcgggat gagcaacccc atccatctgt     960
acttcggtgt gcggagcgag caggacatct atgacgagga acgccttcac gcattggctg    1020
caaggtttcc gaatctcaag gtgaatgtcg ttgttgcaac aggccctgcc ggccctggtc    1080
gtcgatccgg cctggtcacc gatctgatcg gccgtgactt gcccaatttg gcgggatggc    1140
gcgcctacct gtgtggcgct ccggccatgg tcgaggccct gaacctgctc gttgctcgcc    1200
taggcatagt acccgggcac atccatgccg atgcgttcta tcccagcggc gtctgagcga    1260
aggcaccatg cgaacccaat tcaacccaag gataccaagc catgagtgaa ccccaacgat    1320
taaaacccgt gtttccccaa gatccgaaat ggccgggcga aggtagcagc cgcgttccct    1380
tctgggccta cacccgcgaa gacctgtaca agcgcgaatt ggagcgcctg ttctatgcaa    1440
accactggtg ctatgtaggc ctggaagccg agattccgaa tccaggcgac ttcaagcgaa    1500
cggtgatcgg tgagcgctcg gtcatcatgg tgcgtgatcc ggatggcggc atcaacgtgg    1560
tggagaacgt ctgcgcccac cgtggcatgc gcttttgccg cgagcgccac ggcaacgcca    1620
aggacttctt ctgcccctac caccagtgga actacagcct caagggtgac ctgcagggcg    1680
tgcccttccg ccgtggcgtc aagcaggacg gcaaggtcaa cggcggcatg cccaaggact    1740
tcaaactcga agaacacggc ctgaccaagc tcaaggtggc cgcccgaggc ggtgcagtgt    1800
ttgcctcttt tgaccacgat gtcgagccct tcgaggagtt cctgggccca accatcctgc    1860
attacttcga ccgcgtcttc aacggccgca agctcaagat cctgggctac cgccgccagc    1920
gcatcccggg caactggaag ctgatgcagg agaacatcaa ggacccctac cacccgggcc    1980
tgctgcacac ctggttctcg accttcgggc tctggcgcgc cgacaacaag tcggaactga    2040
```

```
agatggacgc caagttccgc cacgccgcaa tgatctccac gcgcggtcag ggcggcaaga   2100
acgaggaggt cgtgtccggc gtggacagct tcaaggaaca gatgaaggtg aacgacccgc   2160
gcctgctcga catcgtgccc gagccctggt ggggcggtcc gactgcggtg atgaccacga   2220
tcttccccag cgtgatcatc cagcagcagg tcaacagcgt atcgacccgc cacatccagc   2280
ccaacggtca cggctccttc gatttcgtct ggacccactt cggcttcgag gacgacaacg   2340
aggagtggac ccagcgccgc ctgatccagg ccaacctgtt cgggccggcg ggcttcgtgt   2400
cggccgatga cggcgaggtg atcgagtggt cgcaggaagg ctttgagcaa aaaccgacgc   2460
accgcaccgt gatcgagatg ggcggtcacg aaatcggcga cacggaccac atggtcaccg   2520
agacgctgat ccgcggcatg tacgactact ggcgcaaggt gatgggggaa taaacatggt   2580
cgacttcaaa acctatttcg aactgctgaa cctgtacagc gactacgcca tggtgtgcga   2640
ctccgccaat tgggagaagt ggcctgattt cttcatcgag accggcacct accgcctgca   2700
gccgcgcgag aactttgagc agggcctgcc gctgtgcctg ctggcgctgg agagcaaggc   2760
aatgatccgc gaccgggtgt acggcgtcaa ggaaaccatg taccacgacc cctactacca   2820
gcgccacatc gtgggcacgc cgcgcgtgct gtcagtggag cgtgacgcgg atggcgagcg   2880
catcaccgcc gaagccagct atgccgtgat tcgcaccaag tacgacggcg attccacgat   2940
tttcaacgcc ggctattacc gagacgtgat cgtgcgcacg cccgagggcc tcaagctgaa   3000
gtcgcgcctg tgcgtttacg acagcgaaat gatccccaac tctgtgatct atccaatctg   3060
aggcgacgca catgactcag aactggattg atgcagcttg tcttgacgac atccctgaag   3120
gcgatgtggt cggcgtcaaa gttaacggca aggaaattgc gctctacgag gtcgaggggg   3180
agatttatgc caccgataac ctgtgcacgc acggcgctgc gcgcatgagc gatggctttc   3240
tcgaaggccg ggaaattgaa tgccctctgc accaaggtcg atttgatgtc tgcacaggca   3300
aagccttgtg cacccccctg acaaaggaca tcaaaaccta ccccgtcaaa attgaaaaca   3360
tgcgcgtgat gctcaaaatg gagtaagact gctttcaatc aggaggctga attcggacca   3420
atcaaccaaa ccatccaatc cacccattag ccaaaaaacg gagacattaa tcatgattta   3480
tgaaaatttg gtgagtgaag cagggctgac gcaaaagcac ctgattcatg gcgacaaaga   3540
acttttccag cacgaattga agaccatctt cgcgcgaaac tggctttttc tgacccatga   3600
cagcttgatt ccctcccccg gcgactatgt cacagccaaa atgggtgtcg atgaagtcat   3660
cgtctcccgc cagaacgatg gctcggtgcg agcctttttg aatgtttgcc gtcaccgggg   3720
caagacacta gttcacgctg aagccggaaa tgcgaaaggc tttgtgtgca gttaccacgg   3780
ctggggcttt ggctccaacg gcgaactgca aagcgttccc tttgaaaaag agttgtacgg   3840
cgatacgatc aaaaagaagt gcctgggctt gaaagaagtc cccgcatcg aaagctttca   3900
tggcttcatt tacggttgtt ttgatgcaga ggcccccacg cttgtcgatt atctgggtga   3960
tgcagcctgg tacttagaac ccatcttcaa gcactctggt ggcctggaac ttgtaggccc   4020
ccccggcaaa gtggtgatca aggccaactg gaaggctcct gcggaaaact ttgtgggtga   4080
cgcgtaccac gttggttgga cgcacgcatc gtctttgcgc tcaggtcagt cgatatttac   4140
ccctcttgcg ggcaacgcta tgcttccacc cgaaggcgcg ggcttacaaa tgaccagcaa   4200
gtatggcagt ggaatgggcg tattgtggga cggctactcc ggtgtccaca gtgctgacct   4260
ggttcccgaa atgatggcat tcggcggcgc aaaacaggaa aaactcgcca aggaaatcgg   4320
cgatgtccgg gcacggattt accgcagcca tctaaactgc acggttttcc cgaacaacag   4380
cattttgacc tgctccggtg tcttcaaggt ctggaacccg atcgatgaaa acacgaccga   4440
```

```
ggtttggacg tatgccatcg tagaaaaaga catgcctgag gacttaaagc gtcgcttggc   4500
tgacgcggtt cagcgcactg tcggaccagc aggattctgg gaaagcgacg acaacgacaa   4560
catggagacg gagtcgcaaa atgccaagaa ataccaatcc agcaacagtg atctgattgc   4620
caatttgggt ttcggcaagg acgtctacgg cgacgaatgc tatccgggcg tcgttgccaa   4680
atcggcaatc ggcgaaacca gctatcgcgg attctaccgt gcctaccagg ctcacatcag   4740
cagctccaat tgggccgagt tcgaaaacac ctcccgaaat tggcacaccg aactcaccaa   4800
gacgactgat cgctaatcca ggagccaacc atgatgatca atacccagga agacaagctg   4860
gtctccgcgc acgacgccga agaatttcat cgtttcttcg tcgggcacga cagcgatctg   4920
cagcaagaag tcaccacact cctgacccga gaagcacatc tgttggacat tcaggcctac   4980
aacgcctggc ttgaacactg cgttgccccc gagatcaaat accaagtgat ctcgcgagaa   5040
tttcggtcca cttccgagcg tcgataccaa ctgaatgatg cggtgaacat ctacaacgag   5100
aactatcagc acctgaaagt tcgagttgag catcagatgg acccgcagaa ctgggccaac   5160
agcccgaaga tccgtttcac ccgcttcgtc accaatgtca cagcggccaa ggacaagatc   5220
gtaccggatc tgctgcatgt gcgttccaac ctcattctcc accgcgccag acgcggcaat   5280
caagttgacg tcttctatgc aacgcgggaa gacaaatgga aacgcatcga aggtggtggc   5340
atccaactgg tggaacgtct tgtggactac ccggagcgca ttctccagac ccacaatctg   5400
atgaccttcc tgtgaaccct ggggatgcct gccttgatgg cggtcatcct tgattgtttt   5460
aaacagaaat ttattgccat gaacatacag caagttattg ccattactgg cgccggttca   5520
ggcatcgggc tggaactggt tcgatccttc aaagcagctg gctattgcgt gtccgcactt   5580
gttcgcaacg aggaacaaga ggcgggcctt cgcagtgaat tcaaagatgc cattgagatc   5640
gtagcgggcg atgtttgtga tcacgccacc aatgaaaagc tggtcaacaa ggcggtcgcc   5700
aggttcggac acctcgactg cttcatcgga aatgccggga tatgggatta catgctgggc   5760
gtcgacgagc catgggagaa attatccggc agtttcgagg agatatttga tatcaacgtc   5820
aaaagctatt tcagcggcat cagcgcggcc ttgccggaac tcaaaaaaac gaacggatcg   5880
gtcgtagtga cggcttccgt ttcttcctat gcggccggcg gcggcggctc ttgctacatt   5940
gccagcaagc atgcagtgct gggtatggtc aaggcgttgg cctacgaatt ggccccgcac   6000
atccgggtca atggcgttgc gccaggtggt acggtcactt ctttggctgg gccggcaagc   6060
gccggtttcg acaaaaccaa aatgaaagac atgcccggca tcgatgacat gatcaagggc   6120
ctgacccccct tggggttcgc agcaaggccc gaggacgtgg tggcaccgta tcttttgctg   6180
gcctcccggg aacaagggaa gttcatcact ggcaccgtaa tcggcattga tggcggcatg   6240
gcgctcggtc gaaagtgaat tttcaatcaa atcagatttt tcaaccccat tcccaggaga   6300
caacccatga agacgaaatt gttcatcaac aacacctgga gcgcttcgag tgacaaaaag   6360
tcattcgatc gcaagcaccc tgtcagtggc gaggtcgtga cccaatgcgc gaacgccacg   6420
gtggacgatg cggtcaatgc ggctcgagcc gctcaagagg cgttcaagtc ctggaaggcc   6480
gtcggaccct cggagcggcg gcgccttctt ttgaaggtgg cagacgtcat ggagagcaaa   6540
acgcccgagt tcatcgaagt gatggccaag gaagtgggag cctccgcgct gtgggcgggg   6600
ttcaacgtgc acctgtcggc caatgtattc cgggaagccg cctcactggc cacccaaatt   6660
caaggcgaaa ccattccgac ggacaagcct gacaccctgt caatgacgct gcgtcacgct   6720
gtcggcccca tcttgagcat cgttccctgg aacggcaccg ccgtgctcgc ggcgcgggcc   6780
atcgcttatc cgctggtctg cggcaatacc gttgtgttca aaggctccga gttcagcccc   6840
```

```
ggtacgcacg ctttgatcac caagtgcctg caggaggccg acctgcctgc tggcgtgctc    6900
aactatctga actcctcccc ggaccggtcg cccgatattg cagatgcgct gatttcgtct    6960
aaagagattc gtcgcatcaa cttcacaggc tccactcgcg tggggcgcat catcgcccag    7020
aaatcggccc agcatctcaa gcgctgcttg ctggagttgg gtggaaagtc cccgctgatc    7080
gttctggacg acgctgacat cgacgcggca gtcaaggccg cggtgttcgg cagcttcctg    7140
ttccaaggcc agatctgcat gtccaccgaa cgcctggtgg tcgacgaaaa gatcgcggac    7200
gaatttgtcg cgaagttcgt cgagaaaacc aagcagttga gcgcaggcga tccatgcgtc    7260
acaggggact gcatcatcgg cccgatggtg tcgcccaact cgggtgaccg aatcaatggt    7320
ctgttcaaag atgccatcag caagggcgcc aaggtcgtgt gcggcggcat cgccgagggt    7380
gcggtcatgc ccgccacgat cttggaccac gtgacagccg acatgcagat ctacgatgag    7440
gaaaccttcg gtcccatcac tgtggttatc cggtgcaaga gcgaagcgga cgccatccgc    7500
attgccaatg acagcgccta cggcctgtca tcgggcgtgt ttggccggga cgtgaaccgg    7560
gctctgcgcg tgggcatggc gatcgaatac ggctcggtcc atatcaacgg ctccaccgta    7620
cagaacgagg ctcaggcgcc ttatggcgga acaaaggcca ccggttatgg ccgcttcgac    7680
ggacgcgcgg tgatcgacga gttcacggaa ctcaagtggc tgaccattga accattcgag    7740
cagcagtatc ccttctaagc tgaagcaaca aaggagttaa accatgaaca agccagcaac    7800
tgtcattgaa ttggggtaca tgggcatttc ggtcaaggat cccgcagcgt ggaaatcctt    7860
tgccgcaaac atgctgggac ttcaagtcct cgatgagggt gacaaggatc gcttctatct    7920
gcgaatggac aattggcacc atcggatcgt ggttcatcac aacggtcaag atgaccttga    7980
atacctgggc tggcgtgtcg ccggtcaacc ggaattcgag gcattgggtc aaaagctcgt    8040
ggacgcaggc tacaaagtcc gcgtgtgcga caaagccgaa gcacaagaac ggatggtgct    8100
gggcctgatg aagacagaag atccgggggg caacccgacc gagattttct ggggaccccg    8160
gattgacctg aacaacccct tccatcccgg tcgtcccttg cacgggaaat ttctaaccgg    8220
tgatcagggc ctgggccact gcatcgtgcg tcagaacgat gttgaagcgg cacgtaagtt    8280
ctatagcttg ctgggatttc gtggagatgt cgagtaccgc cttcctttgc ccaacggcat    8340
gacggctgag ttgacgttca tgcattgcaa tgctcgcgat cattccatcg ctttcggtgc    8400
aatgcctgcg gccaagcgcc tcaatcatct gatgattgaa tacactcata tcgaagattt    8460
gggttgcaca caccagcttt tcacgaagga aaagattgac attgccttgc aattgggcat    8520
ccattccaac gataaggcgc tgacgttcta cggggcaaca ccttccggct ggctgataga    8580
acctgggtgg cgaggcgccc ccgccattgc tgaatcggaa tattacgtcg gcgacatttt    8640
cggccacacc atcgaggcca ccggttatgg attggacgtc aaactgagct agccatgtaa    8700
cagatgcgaa atcgatcgca tctgttttcc ccattcagtt ccatataaaa aaaaggagac    8760
aatgatgatc aaagaagcca tttccctcgc aggacttggg atgctgatgc tcagcaccgc    8820
atatgccgag gattcccgtt ggtcatatcg catcggcgcc accaacgtag ccttcgatgc    8880
gagcgccaaa gtttcgattg acggaacaag ggtgccgggt ggaagcgctg acgccagtga    8940
caacaacgct ttgacatttg acttcggtta catcattaac gataactgga acgcgcgatt    9000
aattgttggc attccaccca ccacaaaagt gacgggcgca ggcacgctgc ctccgatttt    9060
gctgggccgt gtccaatatg ctcctgcagt tttgtcggcg acctacaacc tgccacagat    9120
gggattggtt cgcccgtacg tgggggcggg gatcaactac actcgaattc tgaaaagcaa    9180
agatgccaat ctgacctcat tcgatgcaga tcatgcgtgg gcacttgtgc tccacatcgg    9240
```

```
tgcagaagcg aacatcaacc gcgactggtt cgtcagcttt gatatccgaa aactttatct    9300
aaaaacagat gcatcggggt ttcttgggcc tcaagttgct acagcccgcg taacgttgaa    9360
tccgctgctg acgtcgattg cgattggccg gagattctga tcggtccaca ttgattccaa    9420
aattctgttt gcatcaccat tttcaaagga aatttgaatg acaagaaaga cgagcaaagc    9480
ggtgcgcctg accgccgcgg atattcaagg cgcatgggtc atcatgccga ccccgtccac    9540
gccggatgcc tcggactggc gcagcacgca cacggtcgat ctcgacgaga cggcccggat    9600
tgtcgaggag ttgattgcgg ccggcgtcaa tggcattctg agccacggca cctttggcga    9660
atgcgcgacg ctgacgtggg aggagaagcg ggattttgtt tcaacggtcg tggaaaccgc    9720
gcgcggtcga gtgccctact tctgcggcac aacggccttg aatacccgtg aagtcatacg    9780
ccagacccgc gaattgatcg acattggcgc ccaaggaaca atgctcggcg tgccgatgtg    9840
ggtgaagatg gatctgccta ctgccgtgca attttatcgc gatgtggcgg aagcagtgcc    9900
agatgcagcc atcgctgtct acgccaaccc ggaggctttc aaatttgatt ttcctcgccc    9960
gttttgggcc gaaatgtcca aaatcccgca ggttgtcaca gccaagtact tgggcatcgg   10020
gatgctggac ttggatctga aattggcccc aaatattcgc ttccttccgc atgaggatga   10080
ctactacgct gcggcccgga tcaatcccga gcgcatgact gctttctggt ctagcggttc   10140
catgtgcggc ccagcgaccg ccccttgtgct gcgcgatgag gtggtaaagg ccaaaaatac   10200
aggtgattgg gccaaggcca aggctatttc agatgacatg cgcgcagccg atgccacact   10260
gtttccacgc ggcgatttct cggaattctc aaaatacaac attggcctcg aaaaagcacg   10320
aatggacgag gccggctggc tcaaggcggg gccgtgccgg ccaccctata cgctggttcc   10380
cgacgaatac cttgcaggtg cccgaaaatc aggcaaggcc tgggccgcac tgcataccaa   10440
gtatgccaag gaattgagga aaaccaaaac ggcaaccaac tcgaaaaaga agtaagtcca   10500
ggccctgagt cagacatctc cgatcagcac aacctgctga tctggaggtt ttctggatta   10560
gtgcagtcgg cgagttaaaa atatgccagt acagagtgga agcgccacaa gcgcagggtg   10620
caccgatcgc cctgaagcaa tctcacacta ttgccatctt ctctttgttg ggagtgcatg   10680
atcgtgatgg tcgattttta tttcgatttt ttgagcccat tttcgtatct ggccaaccac   10740
cgtttgtcgg tgctcgccgg gcgttatgga ttctccatcc agtatcacgc cattgatttg   10800
gcgcgagcaa aaacggccat tggcaacatc gggccatcca atcgggacct caaggtcaag   10860
cttgactact taaaggtgga tttgcagcga tgggccgatc tctataggat tccgttggtt   10920
ttccccccta acttcaacag ccgccgggtg aatgccggac tgtattaccc ggcagccagg   10980
gagcgagccg ctgaatatgt tcgccttgtt ttcgattcgg cttgggggaa agggtgggca   11040
ctggatgctg atagcttgct ggctgaggta tgcgacaagc taaactggga tctcggtgaa   11100
tttgaagatt ttttgaacag cgaaaatgcc gccaaggcat acgacgaaga gacgcaggcg   11160
gccattgacc gaaaggtttt cggggttccc accgtgtttt gggatgatca aatgtggtgg   11220
ggaaatgacc gccttttcat gcttgagagc aggttgcaaa aggaaacgca accataaatt   11280
ctcagtgcat cgatttttct gatttttcaa ccacgccttg attttaagga gtctcatgaa   11340
gctttattac agccccggcg catgttcatc gtcgcctcac atcatccttc gtgaaggtgg   11400
atttgacttt cagctagaaa aggttgatct cggcaccaag gtgactgaga ctggtgttga   11460
ttacaagacc gtcaatcccg tgggcagtgt ccccgctttg caaatggatg atgggcaggt   11520
gctcaccgaa gggccggcca tcgttcaata cctcgctgac cgtgtgccag aaaaatgcct   11580
ggcaccggca gctggctcgc tggagcgtta ccgactgatg gaatggctga atttcatctc   11640
```

```
caccgaattg cataaaagct tcggagcgct gttcagcccg gtgtttccac aggatgccaa   11700
gccagtcatc aaggcccaat tggaaagccg tcttgcccat accgagcaga tgctcggtga   11760
caaggtttgg gccatgggga acgacttttc cgtggtggac gcctacttgt ttactgtgct   11820
tggctggggg gcttacgtga atgtggacct ttcgccttgg cccggcctgc aaggctacct   11880
caaccgcgtg gctgaacgtc cagcagtccg ggccacattg tcggctgaag gtttgatctg   11940
atttttttag agaaataaat aaagcatggc cagcctccat gagcggccag ccatgccgcg   12000
gggcagccgt gggcagcggc ctcgatcaaa tttcgatgca cacagattac gagccccttt   12060
ttccgactgc ggtcggattc agaaattgcc gatacttgct ttggtctgat cgaagcgtga   12120
gggatatcgg cttcccctat ccacatgctt gatgaagagg agcgtatcac catgagtcac   12180
gaacttggcc gactggaaga cctgccgcag gactaccgag acgaactcaa acaacttaac   12240
ctggtgccgc tatggcccag cctacgcgcg gtgctgccgc ccaatgtccc gacccgccag   12300
acgcagccga cttactggtc ctaccagacg ctcaagccgc tgctgctcaa ggccggtgag   12360
ctgacccccа ttgagaaggc cgagcgccgc gtgctggtgc tggccaaccc cggccacggc   12420
ctggagaaga tgcaagccag cgccgccata tacctgggca tgcagttgct gctgcccggc   12480
gagtgggcgc ccagccaccg ccacaccccc aacgcggtgc gcatgatcgt ggagggcgag   12540
ggcgcctaca ccaccgtgga tggcgagaag tgccccatga gccggggtga cctcatcctc   12600
acgcccaccg gcctgtggca cgagcacggc cacgacggca acgagcccgt ggtgtggctc   12660
gatgtgctgg acctgccgct ggtgtactac atggaggcca gctaccacat cgacggcgag   12720
cgccagcagg tcgaccccgg ccggggcgac tgcgcctgga cccgtgcggg cgtggtgcca   12780
acccccgtgt tccagcgcag cgacaagcgc tatcctctct tgcgctaccc ttgggccgac   12840
acccgtgccg ccctgctgtc gctggcggcc gaccagcctg agcaggaatg cgtgcaggtc   12900
acctacgtca accccgaaac gggtgacgac gccgagaaca tcctgggctt ctacgccctg   12960
atgctcaagc ccggccagac cctgcgcctg cccgtgcgtt cgcccgccgt ggtgttccac   13020
cagatcgaag gccgcagcga ggcgcgcatc gccgagtcca ccttcgccct gagggaagcc   13080
gatacctgct gcgcccccgg ctacaccgag gtgacgctga aaaacctctc agccgaccag   13140
ccgtccttca tcttcatggc cgacgaatcg cccctgcacc gcaagctggg cgtctttgaa   13200
aaccgcggct gagccgccgc gaagcaccac acaacaggaa agcaagtgag caactacctc   13260
tggaacccgc ctcccgtcca gtctctgccc gttcgtggca agaccgagcg cttccccatc   13320
aaccgcatct tctgcgtcgg ccgcaactac cacgcccacg cggtggaaat gggccgtccg   13380
gtcgacaaaa gcgtcgaaca agcgttctac ttcaccaaat cgccacaaac cctggtggaa   13440
agtggtgcga ccgtggccta cccgccgcgc accagcaact accactacga gatggagctg   13500
gtgctggcga tcggcaagcc cggcttccgc gtcagcgaag accaggcgca cgagttgatt   13560
tacggctacg ccgccggcct ggacatgacc cggcgcgacc tgcaactggt ggcacgcgac   13620
aagggccgcc cctgggacac cggcaaggac atcgaggaag gctcggtctg ctccgagatc   13680
gtgccgatgc aaggcgtggt ggtggagcag ggcgcgatcg ccctggaagt caacggccag   13740
accaaacagt cgtccaacgt ggacaagctg atctggaacg tccgcgagat cattgccgac   13800
ctctccacct actaccactt gcaacccggc gacctgattt acaccggcac gcccgaaggc   13860
gtgggcgccg tggtggcggg tgacaagatc atcggccgtg tggaaggcat tgccgagatc   13920
agcctgaccg tcggcccggc cgagtgagcc tgcacgcgat gaagctgtac aacttctggc   13980
gcagcggcac gtcacaccgg ctgcgcatcg cgctcaacct caagggcgtg ccctacgagt   14040
```

```
acctggccgt gcacctgggc aaggaagagc acctaaagga cgccttcaag gcgctgaacc   14100
cacagcagtt ggtgcccgcg ctggacacgg gcgcgcaagt gctgatccag tcgccggcca   14160
tcatcgaatg gctggaggaa cagtacccca cgcccgcgct gctgccggcc gacgccgacg   14220
gccgccaacg ggtgcgcgcg ctggccgcca tcgtgggctg cgacatccac cccatcaaca   14280
accgccgcat tctggagtac ctgcgcaaga cgttcggcgc cgacgaggcc gccatcaacg   14340
cctggtgtgg cacctggatc agcgccggtt tcgatgccta cgaagcactg ctggcggtcg   14400
acccgaagcg cggccgctac agctttggcg acacacccac gctggccgac tgctacctgg   14460
tgccgcaggt ggaaagcgcc cgccgcttcc aggtggacct gacgccctac cccctgatcc   14520
gcgcagtgga cgcggcctgc ggcgagttgg acgcatttcg gcgcgccgcg ccagctgctc   14580
aacctgattc ggcttgacga aaaatacgcc gccttcgccc taataaacag ttcgtagccc   14640
tgcccagtga caagtatcaa tgagcaatag tgaaaatgta aggggaaaga ttatgaaaca   14700
tattctgacc cgacgagcag ccctgagcac tttgggctcc cttggaatcg gggccttccc   14760
gggcatgagc ttggcgcaat ccggccctgg cgtggccacg atcgtgattg gactcgccgc   14820
aggcggagcg accgatatgg ctgcccgccg gttgtctgag ggaatgcgcg gtgcctatgc   14880
ttccagtgtg ctcgtcgaca accgcacggg ggcgggtggt cgaattgcga tccagcatgt   14940
taaggccgcg gcaccaaacg gcgcgacact tctgctgacg cctgcatcca tgatgacgct   15000
ataccctcac acgtacaaag atct                                          15024
```

```
<210> 12
<211> 4912
<212> DNA
<213> Artificial Sequence

<220>
<223> A modified DNA sequence

<400> 12
gagctcgttg cgcaagcgct tccatcagtg ggggcatgaa gtacatctcg ccgatgtcgg     60
tcattgccaa gttgaaggtg cgcgtgctgg caaatgggtc gaaagagtca cgggtcgtca    120
gtgccgtctg cagcgtgttg agcgcataga ccacgggctc cgcaagatgc agtgcatacg    180
gtgtcggctc catgcctttt gaggtgcgca agaacaaatc gtcctttagc gccgcacgca    240
gccgtttaag tgaattgctg acggcaggct gcgtcagccc cagtttttcg ccggccgtcg    300
atacgctccg gtcgagcagt agctggttga agaccaccag cagattcaag tcgatgtcgc    360
gcagatccat gatgcctcac cattattcat gctggtgatt ttaactatca gacttgatct    420
atagcgctat accgatcgac gcgccagtat cgcagccatt cggagacaac tgaaaaaaga    480
gcttgcatgg aactggtagt agaacccctc aatttgcatc tgaacgcgga gaccggcagc    540
accctgcttg acgtgctcag gtccaacgag gtccccattt cttatagctg catgtcgggc    600
cgctgcggca cttgccgttg ccgtgtgatt gccggccatc ttcgcgataa cggctccgag    660
```

```
acagggcgcc cgcaggcagg aaagggggcc tatgtcctgg cctgtcaggc ggttctgacc      720
gaagactgca cgatcgagat tcctgaatct gacgagatcg tggttcaccc ggcgcgcatc      780
gtcaagggga cggtcacagc gatagacgaa gccacccatg acatccggcg cctgcgcatc      840
aaactggcca aaccgcttga gttcagccct ggccagtacg caacggtgca gttcacgccc      900
gaatgcgtcc gcccctattc gatggccggg ctgcctagcg atgcggaaat ggagtttcag      960
attcgcgcgg ttccgggcgg gcatgtcagc aactacgttt tcaatgaact gtccgtaggc     1020
gcttcggtgc ggatcagcgg ccccctcgga acggcctatc tgcggcgcac gcacaccggc     1080
cccatgcttt gtgtgggggg tggaacaggt ctggcgcccg tcctttcgat cgttcgaggc     1140
gcactggaaa gcgggatgag caaccccatc catctgtact tcggtgtgcg gagcgagcag     1200
gacatctatg acgaggaacg ccttcacgca ttggctgcaa ggtttccgaa tctcaaggtg     1260
aatgtcgttg ttgcaacagg ccctgccggc cctggtcatc gatccggcct ggtcaccgat     1320
ctgatcggcc gtgacttgcc caatttggcg ggatggcgcg cctacctgtg tggcgctccg     1380
gccatggtcg aggccctgaa cctgctcgtt gctcgcctag gcatagtacc cgggcacatc     1440
catgccgatg cgttctatcc cagcggcgtc tgagcgaagg caccatgcga acccaattca     1500
acccaaggat accaagccat gagtgaaccc caacgattaa aacccgtgtt tccccaagat     1560
ccgaaatggc cgggcgaagg tagcagccgc gttcccttct gggcctacac ccgcgaagac     1620
ctgtacaagc gcgaattgga gcgcctgttc tatgcaaacc actggtgcta tgtaggcctg     1680
gaagccgaga ttccgaatcc aggcgacttc aagcgaacgg tgatcggtga gcgctcggtc     1740
atcatggtgc gtgatccgga tggcggcatc aacgtggtgg agaacgtctg cgcccactgt     1800
ggcatgcgct tttgccgcga gcgccacggc aacgccaagg acttcttctg cccctaccac     1860
cagtggaact acagcctcaa gggtgacctg cagggcgtgc ccttccgccg aggcgtcaag     1920
caggacggca aggtcaacgg cggcatgccc aaggacttca aactcgaaga acacggcctg     1980
accaagctca aggtggccgc ccgaggcggt gcagtgtttg cctctttcga ccacgatgtc     2040
gagcctttcg aggactgtgc gtgtacgaca gcgagatgat ccccaactcc ctcatctacc     2100
ccatttgagg ttgcagaaca tgtccgagaa ctggattgat gccatcgcac gggacgctgt     2160
gcctgagggc gatgtggtcg gagtcatcgt ggcaggcaaa gacattgcct tctatgaggt     2220
ggaaggtgag gtcttcgcca ccgacaactt gtgtacccac ggggctgcgc gcttgagcga     2280
cggctttctc gaaggccggg aaattgaatg tcctttgcat caaggccgat tcgatgtttg     2340
cacgggtaaa gccttgtgca cacccctgac acaggacatc aaaacctacc ccgtaaaaat     2400
cgaaaacatg cgcgtgatgc tcaagctgga ctaaaactct ttgcaggagg aaagccaaat     2460
ccggaaatca ccccacccaa cccaatcact acccgttttc aaacaagagg agataagcaa     2520
ttatgagtta ccaaaactta gtgagtgaag cagggctgac gcaaaagctc ctgattcatg     2580
gcgacaaaga acttttccag cacgaattga agaccatctt cgcgcggaac tggctttttc     2640
tgacccatga cagtctgatt ccctcccccg gcgactatgt cacagccaaa atgggcgtcg     2700
atgaagtcat cgtctcccgc cagaacgatg gctcggtgcg agcctttttg aatgtttgcc     2760
gtcaccgggg caagacacta gttcacactg aagccggaaa tgcgaaaggc tttgtgtgcg     2820
gctaccacgg ctggggctac ggttccaacg gcgaactgca aagcgttccc tttgaaaaag     2880
agttgtacgg agatgcgatc aaaaagaaat gcctgggctt gaaagaagtc cccgcatcg      2940
aaagctttca tggctttatc tatggctgtt ttgatgcaga agctcccccg ctcatcgatt     3000
atctgggtga tgcagcctgg tacctggaac ccaccttcaa gcactctggt ggcctggaac     3060
```

```
ttgtaggccc ccccggcaaa gtggtggtta aggccaactg gaagcctttt gcggaaaact   3120
ttgtaggtga catctaccac gttggttgga cgcacgcagc ggctttgcgc gcagggcagt   3180
cggtatttag ttctcttgcg ggcaacgcta agcttccacc cgaaggcgcg ggcttgcaaa   3240
tgaccagcaa gtatggcagt ggaatgggct taacgtggga ctactactcc ggtaacttca   3300
gcgctgatat ggttcccgat ctgatggcat tcggcgccgc aaaacaggaa aaactcgcca   3360
aggaaatcgg cgatgtccgg gcacggattt accgcagcat tctgaacggc acggttttcc   3420
cgaacaacag ctttttgacc ggctccgcta ccttcaaggt ctggaacccg atcgatgaaa   3480
acacgaccga ggtttggacg tatgccttcg tagaaaaaga catgcctgag gacttaaagc   3540
gtcgcttggc tgacgcggct cagcgcagtg tcggaccagc aggattctgg gaaagcgacg   3600
acaacgaaaa catggagacg ttgtcgcaaa atgccaagaa ataccaatcc agcaacagtg   3660
atcagattgc cagtttgggt ttcggcaagg acgtctacgg cgacgaatgc tatccgggcg   3720
tcgttggcaa atcggcaatc ggcgaaacca gctatcgcgg attctaccgt gcctaccagg   3780
ctcacatcag cagctccaat tgggccgagt tcgaaaatgc ctcccgaaat tggcacaccg   3840
aactcaccaa gacaactgat cgctaatcca ggagccaacc atgatgatca atacccagga   3900
agacaagctg gtctccgcgc acgacgccga agaatttcac cgtttcttca tcgtacaaga   3960
tgatgcacta ctgcaagaag tcaacacgct cctgacccgc gaagcgcacc tgctggacat   4020
tcaggcctac aaagcctggc ttgaacactg cgttgccccc gagatcaaat accaagtgat   4080
ctcgcgagaa cttcgctcca cttccgagcg tcgataccaa ctgaatgatg cggtgaatct   4140
ctacaacgag aactatcaac agctgaaagt tcgagttgaa caccagatgg atcctcagaa   4200
ctgggccaac aacccgaaga tccgcttcac ccgcttcgtc accaatgtca cggcggccaa   4260
ggacaagagc gcaccggaaa tgctgcatgt gcggtccaac ctcattctcc atcgcgccag   4320
acgagaaaac caagttgacg tcttctatgc aacgcgtgaa gacaaatgga aacgcatcga   4380
aggtggtggt atcaaattgg tcgaacgctt tgtggactac ccggagcgca ttccccagac   4440
ccacaacctg ttggtcttcc tgtgaaccct ggggatgcct gcctggatgg cgggcattcg   4500
tgattatttt taacagaaat ttattgccat gaacacacag caagttgttg ccatcactgg   4560
cgccggctcg ggcattggtt tcgagttggt tcgctctttt aaggcagccg gttatcgcgt   4620
atctgcactc gttcgcaacg aggagcaaga ggcgggtctt cgcagtgaat tcaaagacga   4680
cattgagatc gtggcgggcg atgtccgtga tcacgccacc aatgagaagc tggttaaaca   4740
gacggttgcc aagttcgggc gcctggattg cttcatcgga aatgccggga tatgggatta   4800
catgctgagc atcgatgatg cttgggagaa attctcgggc agtttcgacg agatatttga   4860
catcaacgtc aaaagctatt tcagcggcat cagcgcggcc ttgccggagc tc           4912
```

<210> 13

<211> 6779

<212> DNA

<213> Artificial Sequence

<220>

<223> A modified DNA sequence

```
<400> 13
atgcatacgt agcggtggcg aaagaggcgc cgctggaaga atccggtctg tagctctggc    60
agaagaccca aggcgagatc aaccgcaccg gactccatat cctccttcag attgccagca   120
ttcgggcgca gcgtgctgat ctggatgtga ggagctcgtt gcgcaagcgc ttccatcagt   180
gggggcatga agtacatctc gccgatgtcg gtcattgcca agttgaaggt gcgcgtgctg   240
gcaaatgggt cgaaagagtc acgggtcgtc agtgccgtct gcagcgtgtt gagcgcatag   300
atcacgggct ccgcaagatg cagtgcatac ggtgtcggct ccatgccttt tgaggtgcgc   360
aagaacaaat cgtcgttcag cgctgtacgc agccgtttaa gtgaattact gacggcaggc   420
tgcgtcagcc ccagtttttc gccggccgtc gatacgctcc ggtcgagcag tagctggttg   480
aagaccacca gcagattcaa gtcgatgtcg cgcagatcca tgacctcacc ctcaccatta   540
ttcatgctgg tgattttaac tatcagactt gatctatagc gctataccga tcgacgcgct   600
agaatcgcag ccattcggag acaactgaaa aaagagcttg catggaactg gtagtagaac   660
ccctcaattt gcatctgaac gcggagaccg gcagcaccct gcttgacgtg ctcaggtcca   720
acgaggtccc catttcttat agctgcatgt cgggccgctg cggcacttgc cgttgccgtg   780
tgattgccgg ccatcttcgc gataacggcc ccgagacagg gcgcccgcag gcaggaaagg   840
gggcctatgt cctggcctgt caggcggttc tgaccgaaga ctgcacgatc gagattcctg   900
aatctgacga gatcgtggtt cacccggcgc gcatcgtcaa ggggacggtc acagcgatag   960
acgaagccac ccatgacatc cggcgcctgc gcatcaaact ggccaaaccg cttgagttca  1020
gccctggcca gtacgcaacg gtgcagttca cgcccgaatg cgtccgcccc tattcgatgg  1080
ccgggctgcc tagcgatgcg gaaatggagt ttcagattcg cgcggttccg ggcgggcatg  1140
tcagcaacta cgttttcaat gaactgtccg taggcgcttc ggtgcggatc agcggccccc  1200
tcggaacggc ctatctgcgg cgcacgcaca ccggccccat gctttgtgtg gggggtggaa  1260
caggtctggc gcccgtcctt tcgatcgttc gaggcgcact ggaaagcggg atgagctacc  1320
ccatccatct gtacttcggt gtgcggagcg agcaggacat ctatgacgag gaacgccttc  1380
acgcattggc tgcaaggttt ccgaatctca aggtgaatgt cgttgttgca acaggccctg  1440
ccggccctgg tcatcgatcc ggcctggtca ccgatctgat cggccgtgac ttgcccaatt  1500
tggcgggatg gcgactgcat cctgtgtggc gctccggcca tggtcgaggc cctgaacctg  1560
ctcgttgctc gcctaggcat agtacccggg cacatccatg ccgatgcgtt ctatcccagc  1620
ggcgtctgag cgaaggcacc atgcgaaccc aattcaaccc aaggatacca agccatgagt  1680
gaaccccaac gattaaaacc cgtgtttccc caagatccga aatggccggg cgaaggtagc  1740
agccgcgttc ccttctgggc ctacacccgc gaagacctgt acaagcgcga attggagcgc  1800
ctgttctatg caaaccactg gtgctatgta ggcctggaag ccgagattcc gaatccaggc  1860
gacttcaagc gaacggtgat cggtgagcgc tcggtcatca tggtgcgtga tccggatggc  1920
ggcatcaacg tggtggagaa cgtctgcgcc caccgtggca tgcgcttttg ccgcgagcgc  1980
cacggcaacg ccaaggactt cttctgcccc taccaccagt ggaactacag cctcaagggt  2040
gacctgcagg gcgtgccctt ccgccgtggc gtcaagcagg acggcaaggt caacggcggc  2100
atgcccaagg acttcaaact cgaagaacac ggcctgacca agctcaaggt ggccgccccg  2160
aggcggtgca gtgtttgcct ctttcgacca cgatgtcgag cctttcgagg acttcctggg  2220
ccaaccatcc tgcactactt cgatcgcgtc ttcaatggcc gcaagctcaa gatcctgggc  2280
taccgccgcc agcgcatccc gggcaactgg aagctgatgc aggagaacat caaggacccc  2340
```

```
taccacccgg gcctgctgca cacctggttc tcgaccttcg ggctctggcg cgccgacaac    2400
aagtcggaac tgaagatgga cgccaagttc cgccacgccg caatgatctc cacgcgcggt    2460
cagggcggca agaacgagga ggtcgtgtcc ggcgtggaca gcttcaagga acagatgaag    2520
gtgaacgacc cgcgcctgct cgacatcgtg cccgagccct ggtggggcgg tccgactgcg    2580
gtgatgacca cgatcttccc cagcgtgatc atccagcagc aggtcaacag cgtatcgacc    2640
cgccacatcc agcccaacgg tcacggctcc ttcgatttcg tctggaccca cttcggcttc    2700
gaggacgaca acgaggagtg gacccagcgc cgcctgatcc aggccaacct gttcgggccg    2760
gcgggcttcg tgtcggccga tgacggcgag gtgatcgagt ggtcgcagga aggctttgag    2820
caaaaaccga cgcaccgcac cgtgatcgag atgggcggtc acgaaatcgg cgacacggac    2880
cacatggtca ccgagacgct gatccgcggc atgtacgact actggcgcaa ggtgatgggg    2940
gaataaacat ggtagacttc aaaacctatt tcgaactgct gaacctgtac agcgactacg    3000
ccatggtgtg cgactccgcc aattgggaga agtggcctga tttcttcatc gagaccggca    3060
cctaccgcct gcaaccgcgc gaaaacttcg agcaggactt gccgctgtgt ctgctggcgc    3120
tggagagcaa ggccatgatt cgtgaccgag tgtacggtgt caaggaaacc atgtaccacg    3180
atccctacta ccagcgccac atcgtaggca cgccgcgcgt gctgtcagtg gagcgtgatg    3240
cggacggcga gcgcatcacc gccgaagcca gctatgccgt gattcgcacc aagtacgacg    3300
gcgattccac gattttcaac gccggctatt accgagacgt gatcgtgcgc acgcccgagg    3360
gcctcaagct gaagtcgcgc ctgtgcgtgt acgacagcga aatgattccc aactccatca    3420
tctaccctat ctgagaagga atccaatgag cgagaactgg atcgacgccg ccgcccgcga    3480
cgaggtgcca cgagggcgac gtgatcggca tcaatatcgt cggcaaggag attgcctcta    3540
cgaggtggcg ggcgagatct acgccaccga caacacctgc actcacggcg ccgcccgcat    3600
gagcgatggc tttctcgaag gccgggaaat tgaatgtcct ttgcatcaag gccgattcga    3660
tgtttgcacg ggtaaagcct tgtgcacacc cctgacacag gacatcaaaa cctaccccgt    3720
aaaaatcgaa aacatgcgcg tgatgctcaa gctggactaa atgctcaagc tggactaaaa    3780
ctctttgcag gaggaaagcc aaatccggaa atcaccccac ccaacccaat cactacccgt    3840
tttcaaacaa gatgagacaa gcaattatga gttaccaaaa cttagtgagt gaagcagggc    3900
tgacgcaaaa gcacctgatt tatggcgaca aagaactttt ccagcacgaa ttgaagacca    3960
tcttcgcgcg gaactggctt tttctgaccc atgacagtct gattccctcc cccggcgact    4020
atgtcaaagc caaaatgggc gtcgatgaag tcatcgtctc ccgccagaac gatggctcgg    4080
tgcgagcctt tttgaatgtt tgccgtcacc ggggcaagac aatagttgac gctgaagccg    4140
gaaatgcgaa aggctttgtg tgcggttacc acggctgggg ctatggctcc aacggcgaac    4200
tgcaaagcgt tccctttgaa aaagagttgt acggagatgc gatcaaaaag aaatgcctgg    4260
gcttgaaaga agtcccccgc atcgaaagct ttcatggctt tatctatggc tgttttgatg    4320
cagaagctcc cccgctcatc gattatctgg gtgatgtagc ctggtacctg gaacccacct    4380
tcaagcactc tggtggcctg gaacttgtag gcccccccgc caaagtggtg gttaagggca    4440
actggaaggt tttgcggaa aactttgtag gtgacatcta ccacattggt tggacgcacg    4500
catctatttt gcgcgcaggg caggcgatat ttgctcctct tgcgggcaac gctatgcttc    4560
cacccgaagg cacgggcttg caagcgacca ccaagtatgg cagtggaatt ggcgtatcgt    4620
tggacgccta ctccggtgtc cagagcgctg atctggttcc cgaaatgatg gcattcggcg    4680
gcgcaaaaca ggaaaagctc gccaaagaaa tcggcgatgt ccgggcgcgg atttaccgca    4740
```

```
gccaagtgaa cggcacggtt ttcccgaaca actgcttttt gaccggcgcc ggtgtcttca   4800
aggtctttaa cccgatcgat gaaaacacga ccgaggcttg gacgtatgcc atcgtagaaa   4860
aagacatgcc tgaggactta aagcgtcgct tggctgacgc ggctcagcgc tctgtcggac   4920
cagcaggata ctgggaaagc gacgacaacg acaacatggt gttgtcgcaa aatgccaaga   4980
aataccaatc cagcaacagt gatctgattg ccgatttggg tttcggcaag gacgtctacg   5040
gcgacgaatg ctatccgggc gtcgttagca aatcggcatt cagcgaaacc aaccatcgcg   5100
gattctaccg tgcctaccag gctcacatca gcagctccaa ttgggccgag ttcgaaaaca   5160
cctccccgaaa ttggcacacc gaactcacca agacgactga tcgctaatcc aggagccaat   5220
catgatgatc aatacccagg aagacaagct ggtctccgcg cacgacgccg aagaatttca   5280
ccgtttcttc gtcgggcacg acagcgatct gcagcaagaa gtcaccacac tcctgacccg   5340
cgaagccgac ctgctggaca ttcaggccta caaagcctgg cttgaacact gcgttgcccc   5400
cgagatcaaa taccaagtga tctcgcgaga acttcgctcc acttccgagc gtcgatacca   5460
actgaatgat gcggtgaata tctacaacga gaactatcaa cagctgaaag ttcgagttga   5520
acaccagatg gatcctcaga actggtacaa cagcccgaag atccgcttca cccgcttcgt   5580
caccaatgtc acggcggcca aggacaagag cgcaccggaa atgctgcatg tgcggtccaa   5640
cctcattctc catcgcgcca gacgaggaaa ccaagttgac gtcttctatg caacgcgaga   5700
agacaaatgg aaacgcatcg aaggtggtgg catcaaattg gtcgaacgct ttgtggacta   5760
cccggagcgc agtccccaaa cccacaacct gatgatcttc ctgtgagccc tggggatgcc   5820
tgcctggatg gcgggcattc gtgattattt ttaacagggc ggattcaaaa gtgaagtgca   5880
acaccctgga tttcagtgaa tgagagtgga gtgctgcggg ctattcacaa gcagttcgcg   5940
gtagaccgcc agcggtgatc gtactcccag ccctttcctt ggccggttgt tgatttcatc   6000
ggcaatcgca tcgagctgct cctggctgta gatgctcaga tccgtccctt tgggcaggta   6060
ctggcgcacc aggccattca tgttctcgtt ggagcctctt tgccaagggc tgtgcgggtc   6120
gcagaagtac acggctatgc cggtttgctc actgagcttc ttgtgcatgg acatctcccg   6180
gccctggtcg tacgtcatgc tcaaccgcat cggctgcgca atgcccagca gcttgtccgt   6240
gaaggcctgc aagacgttgg cagcactggc cggcttgaac tcaggcagct tgaccagcat   6300
caccagccgg ctggtgcgtt ccaccaaggt cccgactgcg ctggcattgc cttctccctt   6360
gatgagatca ccttcccagt gcccgggaaa ctggcggtct tcgatctcgg gcgggcgtac   6420
atggatgctc agcatgtcgg gaatctggcc tctgcggtcc ttgcccttgc tgcgcggcag   6480
ccgcttgttg tgggcgtgac gcagggtggc gatcagctct ttgcgcagct ctcccacggg   6540
catggcatag atgcagttgt agatggtttc gtgtgacacg cggtattcat ggcccagggg   6600
atacaaacgt gccagtgtca gggcaatctg ctcgggcgac cagcgcagcc gcagcagatg   6660
gatcaccagg caggacaaga tcgactcggg atgcagcttg atagctggtc tgccacagcg   6720
tctgcgctgc agggactggc actgggcatg tgcgctgcca tagccagatg aagatgcat    6779
```

<210> 14

<211> 449

<212> PRT

<213> Artificial Sequence

<220>
<223> A polypeptide encoded by SEQ ID NO:3

<400> 14

Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
1 5 10 15

Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
20 25 30

Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
35 40 45

Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
50 55 60

Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65 70 75 80

Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
85 90 95

Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
100 105 110

Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
115 120 125

Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
130 135 140

His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
145 150 155 160

Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
165 170 175

Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
180 185 190

Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
195 200 205

Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
210 215 220

Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225 230 235 240

Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
245 250 255

Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
260 265 270

09843250 .042601

Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
275 280 285

Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
290 295 300

Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305 310 315 320

Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
325 330 335

Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val
340 345 350

Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
355 360 365

Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
370 375 380

Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
385 390 395 400

Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
405 410 415

Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
420 425 430

Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
435 440 445

Arg

<210> 15
<211> 449
<212> PRT
<213> Artificial Sequence

<220>
<223> A polypeptide encoded by SEQ ID NO:4

<400> 15

Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
1 5 10 15

Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
20 25 30

Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
35 40 45

```
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
    50                  55                  60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65                  70                  75                  80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
        115                 120                 125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
145                 150                 155                 160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
    210                 215                 220
Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
                245                 250                 255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
            260                 265                 270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
        275                 280                 285
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
    290                 295                 300
Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305                 310                 315                 320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
                325                 330                 335
Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val
            340                 345                 350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
        355                 360                 365
```

```
Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
    370                 375                 380
Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
385                 390                 395                 400
Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
                405                 410                 415
Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
            420                 425                 430
Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
        435                 440                 445
Arg



      <210> 16
      <211> 449
      <212> PRT
      <213> Artificial Sequence

      <220>
      <223> A polypeptide encoded by SEQ ID NO:5

      <400> 16
Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
 1              5                   10                  15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
            20                  25                  30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
        35                  40                  45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
    50                  55                  60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65                  70                  75                  80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
        115                 120                 125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
```

His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
145 150 155 160

Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
165 170 175

Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
180 185 190

Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
195 200 205

Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
210 215 220

Cys Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225 230 235 240

Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
245 250 255

Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
260 265 270

Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
275 280 285

Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
290 295 300

Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305 310 315 320

Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
325 330 335

Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val
340 345 350

Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
355 360 365

Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
370 375 380

Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
385 390 395 400

Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
405 410 415

Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
420 425 430

Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
435 440 445

Arg

```
<210> 17
<211> 449
<212> PRT
<213> Artificial Sequence

<220>
<223> A polypeptide encoded by SEQ ID NO:6

<400> 17
Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Phe Gly Leu Ser Gln
1               5                   10                  15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
            20                  25                  30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
        35                  40                  45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
    50                  55                  60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65                  70                  75                  80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
        115                 120                 125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
145                 150                 155                 160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
    210                 215                 220
Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
```

```
                    245                     250                     255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
                260                     265                     270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
            275                     280                     285
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
        290                     295                     300
Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305                     310                     315                     320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
                    325                     330                     335
Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val
                340                     345                     350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
            355                     360                     365
Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
        370                     375                     380
Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
385                     390                     395                     400
Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
                    405                     410                     415
Tyr Arg Ala Tyr Arg Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
                420                     425                     430
Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
            435                     440                     445
Arg


        <210> 18
        <211> 449
        <212> PRT
        <213> Artificial Sequence

        <220>
        <223> A polypeptide encoded by SEQ ID NO:7

        <400> 18
Met Asn Tyr Lys Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Thr Gln
 1               5                  10                  15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Arg
```

```
            20                  25                  30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
        35                  40                  45
Pro Ser Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
    50                  55                  60
Ile Val Ser Arg Gln Ser Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65                  70                  75                  80
Cys Arg His Arg Gly Lys Thr Leu Val Asn Ala Glu Ala Gly Asn Ala
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Glu Ser Leu
        115                 120                 125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
145                 150                 155                 160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Ile Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
    210                 215                 220
Ala Ser Leu Ala Gly Asn Ala Val Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
                245                 250                 255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
            260                 265                 270
Gly Gly Ser Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
        275                 280                 285
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
    290                 295                 300
Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305                 310                 315                 320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
                325                 330                 335
Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val
```

```
65                  70
```

```
145                 150                 155                 160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
    210                 215                 220
Cys Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
                245                 250                 255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
            260                 265                 270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
        275                 280                 285
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
    290                 295                 300
Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305                 310                 315                 320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
                325                 330                 335
Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val
            340                 345                 350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
        355                 360                 365
Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
    370                 375                 380
Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
385                 390                 395                 400
Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
                405                 410                 415
Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
            420                 425                 430
Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
        435                 440                 445
Arg


      <210> 17
      <211> 449
      <212> PRT
```

```
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Glu Ser Leu
        115                 120                 125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Ser Leu Met
145                 150                 155                 160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Ile Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Thr Gly Glu Ser Ile Phe
    210                 215                 220
Ser Ser Leu Ala Gly Asn Ala Val Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
                245                 250                 255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
            260                 265                 270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Pro
        275                 280                 285
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
    290                 295                 300
Ser Val Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305                 310                 315                 320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
                325                 330                 335
Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ala Val Gln Arg Thr Val
            340                 345                 350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
        355                 360                 365
Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Ile
    370                 375                 380
Ser Asn Leu Gly Phe Gly Lys Asp Val Tyr Gly Asp Ala Val Tyr Pro
385                 390                 395                 400
Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
```

```
                405                 410                 415
Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
            420                 425                 430
Glu Asp Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
        435                 440                 445
Arg



        <210> 20
        <211> 449
        <212> PRT
        <213> Artificial Sequence

        <220>
        <223> A polypeptide encoded by SEQ ID NO:9

        <400> 20
Met Asn Tyr Lys Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Thr Gln
 1               5                  10                  15
Lys His Leu Ile His Gly Gly Glu Gly Leu Phe Gln His Glu Leu Arg
            20                  25                  30
Ala Val Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
        35                  40                  45
Pro Ser Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
    50                  55                  60
Ile Val Ser Arg Gln Ser Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65                  70                  75                  80
Cys Arg His Arg Gly Lys Thr Leu Val Asn Ala Glu Ala Gly Asn Ala
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Glu Ser Leu
        115                 120                 125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
His Gly Phe Ile Tyr Ala Cys Ile Asp Gln Glu Ala Pro Ser Leu Met
145                 150                 155                 160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Ile Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
```

```
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Cys Thr Gly Glu Ser Ile Phe
    210                 215                 220
Ser Ser Leu Ala Gly Asn Ala Val Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
                245                 250                 255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
            260                 265                 270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
        275                 280                 285
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
    290                 295                 300
Ser Val Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305                 310                 315                 320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
                325                 330                 335
Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ala Val Gln Arg Thr Val
            340                 345                 350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
        355                 360                 365
Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Ile
    370                 375                 380
Ser Asn Leu Gly Phe Gly Lys Asp Val Tyr Gly Asp Ala Val Tyr Pro
385                 390                 395                 400
Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
                405                 410                 415
Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
            420                 425                 430
Glu Asp Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
        435                 440                 445
Arg
```

```
<210> 21
<211> 449
<212> PRT
<213> Artificial Sequence
```

```
<220>
<223> A polypeptide encoded by SEQ ID NO:10

<400> 21
Met Asn Tyr Lys Asn Lys Asn Leu Val Ser Glu Ser Gly Leu Thr Gln
 1              5                   10                  15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln Arg Glu Leu Glu
            20                  25                  30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
        35                  40                  45
Pro Ser Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Val Asp Glu Val
    50                  55                  60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65                  70                  75                  80
Cys Arg His Arg Gly Lys Thr Leu Val His Ala Glu Ala Gly Asn Ala
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ala Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Glu Ala Leu
        115                 120                 125
Asp Lys Lys Cys Met Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
His Gly Phe Ile Tyr Gly Cys Phe Asp Glu Glu Ala Pro Ser Leu Lys
145                 150                 155                 160
Asp Tyr Met Gly Asp Ala Gly Trp Tyr Leu Glu Pro Met Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Ile Gly Pro Pro Gly Lys Val Ile Ile Lys
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Thr Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Gln Ser Val Phe
    210                 215                 220
Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
                245                 250                 255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
            260                 265                 270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Glu Val Arg
        275                 280                 285
```

```
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
    290                 295                 300
Ser Phe Leu Thr Cys Ser Gly Val Phe Lys Val Trp His Pro Ile Asp
305                 310                 315                 320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Met Val Glu Lys Asp Met
                325                 330                 335
Pro Glu Asp Leu Lys Arg Arg Leu Val Asp Ala Val Gln Arg Thr Val
            340                 345                 350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
        355                 360                 365
Val Ser Gln Asn Ala Lys Lys Tyr Gln Ser Arg Asp Gly Asp Leu Val
    370                 375                 380
Ser Asn Leu Gly Phe Gly Gly Asp Val Tyr Gly Asp Glu Val Tyr Pro
385                 390                 395                 400
Gly Ile Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
                405                 410                 415
Tyr Arg Ala Tyr Gly Ala His Ile Ser Ser Ser Ser Trp Ala Glu Phe
            420                 425                 430
Glu Asp Val Ser Lys Asn Trp His Thr Glu Leu Ala Lys Thr Thr Asp
        435                 440                 445
Arg



      <210> 22
      <211> 447
      <212> PRT
      <213> Artificial Sequence

      <220>
      <223> A polypeptide encoded by SEQ ID NO:11

      <400> 22
Met Ile Tyr Glu Asn Leu Val Ser Glu Ala Gly Leu Thr Gln Lys His
 1               5                  10                  15
Leu Ile His Gly Asp Lys Glu Leu Phe Gln His Glu Leu Lys Thr Ile
            20                  25                  30
Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile Pro Ser
        35                  40                  45
Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Val Asp Glu Val Ile Val
    50                  55                  60
```

```
Ser Arg Gln Asn Asp Gly Ser Val Arg Ala Phe Leu Asn Val Cys Arg
65                  70                  75                  80
His Arg Gly Lys Thr Leu Val His Ala Glu Ala Gly Asn Ala Lys Gly
                85                  90                  95
Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly Glu Leu
            100                 105                 110
Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Asp Thr Ile Lys Lys
        115                 120                 125
Lys Cys Leu Gly Leu Lys Glu Val Pro Arg Ile Glu Ser Phe His Gly
    130                 135                 140
Phe Ile Tyr Gly Cys Phe Asp Ala Glu Ala Pro Thr Leu Val Asp Tyr
145                 150                 155                 160
Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Ile Phe Lys His Ser Gly
                165                 170                 175
Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys Ala Asn
            180                 185                 190
Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His Val Gly
        195                 200                 205
Trp Thr His Ala Ser Ser Leu Arg Ser Gly Gln Ser Ile Phe Thr Pro
    210                 215                 220
Leu Ala Gly Asn Ala Met Leu Pro Pro Glu Gly Ala Gly Leu Gln Met
225                 230                 235                 240
Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly Tyr Ser
                245                 250                 255
Gly Val His Ser Ala Asp Leu Val Pro Glu Met Met Ala Phe Gly Gly
            260                 265                 270
Ala Lys Gln Glu Lys Leu Ala Lys Glu Ile Gly Asp Val Arg Ala Arg
        275                 280                 285
Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn Ser Ile
    290                 295                 300
Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp Glu Asn
305                 310                 315                 320
Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met Pro Glu
                325                 330                 335
Asp Leu Lys Arg Arg Leu Ala Asp Ala Val Gln Arg Thr Val Gly Pro
            340                 345                 350
Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr Glu Ser
        355                 360                 365
Gln Asn Ala Lys Lys Tyr Gln Ser Ser Asn Ser Asp Leu Ile Ala Asn
    370                 375                 380
```

Leu Gly Phe Gly Lys Asp Val Tyr Gly Asp Glu Cys Tyr Pro Gly Val
385 390 395 400
Val Ala Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe Tyr Arg
405 410 415
Ala Tyr Gln Ala His Ile Ser Ser Ser Asn Trp Ala Glu Phe Glu Asn
420 425 430
Thr Ser Arg Asn Trp His Thr Glu Leu Thr Lys Thr Thr Asp Arg
435 440 445

<210> 23
<211> 447
<212> PRT
<213> Artificial Sequence

<220>
<223> A polypeptide encoded by SEQ ID NO:12

<400> 23
Met Ser Tyr Gln Asn Leu Val Ser Glu Ala Gly Leu Thr Gln Lys Leu
1 5 10 15
Leu Ile His Gly Asp Lys Glu Leu Phe Gln His Glu Leu Lys Thr Ile
20 25 30
Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile Pro Ser
35 40 45
Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Val Asp Glu Val Ile Val
50 55 60
Ser Arg Gln Asn Asp Gly Ser Val Arg Ala Phe Leu Asn Val Cys Arg
65 70 75 80
His Arg Gly Lys Thr Leu Val His Thr Glu Ala Gly Asn Ala Lys Gly
85 90 95
Phe Val Cys Gly Tyr His Gly Trp Gly Tyr Gly Ser Asn Gly Glu Leu
100 105 110
Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Asp Ala Ile Lys Lys
115 120 125
Lys Cys Leu Gly Leu Lys Glu Val Pro Arg Ile Glu Ser Phe His Gly
130 135 140
Phe Ile Tyr Gly Cys Phe Asp Ala Glu Ala Pro Pro Leu Ile Asp Tyr
145 150 155 160
Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Thr Phe Lys His Ser Gly
165 170 175

```
Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Val Lys Ala Asn
            180                 185                 190
Trp Lys Pro Phe Ala Glu Asn Phe Val Gly Asp Ile Tyr His Val Gly
        195                 200                 205
Trp Thr His Ala Ala Ala Leu Arg Ala Gly Gln Ser Val Phe Ser Ser
    210                 215                 220
Leu Ala Gly Asn Ala Lys Leu Pro Pro Glu Gly Ala Gly Leu Gln Met
225                 230                 235                 240
Thr Ser Lys Tyr Gly Ser Gly Met Gly Leu Thr Trp Asp Tyr Tyr Ser
                245                 250                 255
Gly Asn Phe Ser Ala Asp Met Val Pro Asp Leu Met Ala Phe Gly Ala
            260                 265                 270
Ala Lys Gln Glu Lys Leu Ala Lys Glu Ile Gly Asp Val Arg Ala Arg
        275                 280                 285
Ile Tyr Arg Ser Ile Leu Asn Gly Thr Val Phe Pro Asn Asn Ser Phe
    290                 295                 300
Leu Thr Gly Ser Ala Thr Phe Lys Val Trp Asn Pro Ile Asp Glu Asn
305                 310                 315                 320
Thr Thr Glu Val Trp Thr Tyr Ala Phe Val Glu Lys Asp Met Pro Glu
                325                 330                 335
Asp Leu Lys Arg Arg Leu Ala Asp Ala Ala Gln Arg Ser Val Gly Pro
            340                 345                 350
Ala Gly Phe Trp Glu Ser Asp Asp Asn Glu Asn Met Glu Thr Leu Ser
        355                 360                 365
Gln Asn Ala Lys Lys Tyr Gln Ser Ser Asn Ser Asp Gln Ile Ala Ser
    370                 375                 380
Leu Gly Phe Gly Lys Asp Val Tyr Gly Asp Glu Cys Tyr Pro Gly Val
385                 390                 395                 400
Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe Tyr Arg
                405                 410                 415
Ala Tyr Gln Ala His Ile Ser Ser Ser Asn Trp Ala Glu Phe Glu Asn
            420                 425                 430
Ala Ser Arg Asn Trp His Thr Glu Leu Thr Lys Thr Thr Asp Arg
        435                 440                 445
```

```
<210> 24
<211> 451
<212> PRT
<213> Artificial Sequence
```

```
<220>
<223> A polypeptide encoded by SEQ ID NO:13

<400> 24
Met Arg Gln Ala Ile Met Ser Tyr Gln Asn Leu Val Ser Glu Ala Gly
 1               5                  10                  15
Leu Thr Gln Lys His Leu Ile Tyr Gly Asp Lys Glu Leu Phe Gln His
            20                  25                  30
Glu Leu Lys Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp
        35                  40                  45
Ser Leu Ile Pro Ser Pro Gly Asp Tyr Val Lys Ala Lys Met Gly Val
    50                  55                  60
Asp Glu Val Ile Val Ser Arg Gln Asn Asp Gly Ser Val Arg Ala Phe
65                  70                  75                  80
Leu Asn Val Cys Arg His Arg Gly Lys Thr Ile Val Asp Ala Glu Ala
                85                  90                  95
Gly Asn Ala Lys Gly Phe Val Cys Gly Tyr His Gly Trp Gly Tyr Gly
            100                 105                 110
Ser Asn Gly Glu Leu Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly
        115                 120                 125
Asp Ala Ile Lys Lys Lys Cys Leu Gly Leu Lys Glu Val Pro Arg Ile
    130                 135                 140
Glu Ser Phe His Gly Phe Ile Tyr Gly Cys Phe Asp Ala Glu Ala Pro
145                 150                 155                 160
Pro Leu Ile Asp Tyr Leu Gly Asp Val Ala Trp Tyr Leu Glu Pro Thr
                165                 170                 175
Phe Lys His Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Ala Lys Val
            180                 185                 190
Val Val Lys Gly Asn Trp Lys Val Phe Ala Glu Asn Phe Val Gly Asp
        195                 200                 205
Ile Tyr His Ile Gly Trp Thr His Ala Ser Ile Leu Arg Ala Gly Gln
    210                 215                 220
Ala Ile Phe Ala Pro Leu Ala Gly Asn Ala Met Leu Pro Pro Glu Gly
225                 230                 235                 240
Thr Gly Leu Gln Ala Thr Thr Lys Tyr Gly Ser Gly Ile Gly Val Ser
                245                 250                 255
Leu Asp Ala Tyr Ser Gly Val Gln Ser Ala Asp Leu Val Pro Glu Met
            260                 265                 270
Met Ala Phe Gly Gly Ala Lys Gln Glu Lys Leu Ala Lys Glu Ile Gly
        275                 280                 285
```

```
Asp Val Arg Ala Arg Ile Tyr Arg Ser Gln Val Asn Gly Thr Val Phe
    290                 295                 300
Pro Asn Asn Cys Phe Leu Thr Gly Ala Gly Val Phe Lys Val Phe Asn
305                 310                 315                 320
Pro Ile Asp Glu Asn Thr Thr Glu Ala Trp Thr Tyr Ala Ile Val Glu
                325                 330                 335
Lys Asp Met Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ala Ala Gln
            340                 345                 350
Arg Ser Val Gly Pro Ala Gly Tyr Trp Glu Ser Asp Asp Asn Asp Asn
        355                 360                 365
Met Val Leu Ser Gln Asn Ala Lys Lys Tyr Gln Ser Ser Asn Ser Asp
    370                 375                 380
Leu Ile Ala Asp Leu Gly Phe Gly Lys Asp Val Tyr Gly Asp Glu Cys
385                 390                 395                 400
Tyr Pro Gly Val Val Ser Lys Ser Ala Phe Ser Glu Thr Asn His Arg
                405                 410                 415
Gly Phe Tyr Arg Ala Tyr Gln Ala His Ile Ser Ser Ser Asn Trp Ala
            420                 425                 430
Glu Phe Glu Asn Thr Ser Arg Asn Trp His Thr Glu Leu Thr Lys Thr
        435                 440                 445
Thr Asp Arg
    450


      <210> 25
      <211> 2265
      <212> DNA
      <213> Pseudomonas sp.

      <400> 25
gagggtagag aaatcgaatg ccccttgcat caaggtcggt ttgacgtttg cacaggcaaa      60
gccctgtgcg cacccgtgac acagaacatc aaaacatatc cagtcaagat tgagaacctg     120
cgcgtaatga ttgatttgag ctaagaattt taacaggagg cacccccggc cctagagcgt     180
aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt     240
aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca     300
tgaactgaaa accatttttg cgcggaactg gcttttttctc actcatgata gcctgattcc     360
tgccccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccggca     420
gaacgacggt tcgattcgtg cttttctgaa cgtttgccgg catcgtggca agacgctggt     480
gagcgtggaa gccggcaatg ccaaaggttt tgtttgcagc tatcacggct ggggcttcgg     540
ctccaacggt gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa     600
taaaaaatgt ctggggttga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta     660
```

```
cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggta      720
cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggt      780
tgtgatcaag gccaactgga aggcacccgc ggaaaacttt gtgggagatg cataccacgt      840
gggttggacg cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg      900
caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg      960
catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt     1020
gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc     1080
tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg     1140
ctcgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggaccta     1200
cgccattgtc gaaaaagaca tgcctgagga tctcaagcgc cgcttggccg actctgttca     1260
gcgaacgttc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc     1320
ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt     1380
cggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg     1440
cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cacgtcagca gctccaactg     1500
ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg     1560
ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac     1620
gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc     1680
actacgctgc tgacccagga agcgcatttg ttggacattc aggcttaccg tgcttggtta     1740
gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct     1800
tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa     1860
ctgaaagttc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg     1920
cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta     1980
cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc     2040
ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc     2100
cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg     2160
tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg     2220
aatgtacgtg tatgggcaat caacaagtcg tttcgataac cggtg                     2265


        <210> 26
        <211> 449
        <212> PRT
        <213> Pseudomonas sp.

        <400> 26
Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
1               5                   10                  15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
            20                  25                  30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
        35                  40                  45
```

```
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
    50                  55                  60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65                  70                  75                  80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
        115                 120                 125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
145                 150                 155                 160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
    210                 215                 220
Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
                245                 250                 255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
            260                 265                 270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
        275                 280                 285
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
    290                 295                 300
Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305                 310                 315                 320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
                325                 330                 335
Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Phe
            340                 345                 350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
        355                 360                 365
```

```
Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
    370                 375                 380
Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
385                 390                 395                 400
Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
                405                 410                 415
Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
            420                 425                 430
Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
        435                 440                 445
Arg


      <210> 27
      <211> 2265
      <212> DNA
      <213> Artificial Sequence

      <220>
      <223> A modified DNA sequence

      <400> 27
gagggtagag aaatcgaatg ccccttgcat caaggtcggt ttgacgtttg cacaggcaaa      60
gccctgtgcg cacccgtgac acagaacatc aaaacatatc cagtcaagat tgagaacctg     120
cgcgtaatga ttgatttgag ctaagaattt taacaggagg caccccgggc cctagagcgt     180
aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt     240
aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca     300
tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc     360
tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccggca     420
gaacgacggt tcgattcgtg cttttctgaa cgtttgccgg catcgtggca agacgctggt     480
gagcgtggaa gccggcaatg ccaaaggttt tgtttgcagc tatcacggct ggggcttcgg     540
ctccaacggt gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa     600
taaaaaatgt ctggggttga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta     660
cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggta     720
cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggt     780
tgtgatcaag gccaactgga aggcacccgc ggaaaacttt gtgggagatg cataccacgt     840
gggttggacg cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg     900
caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg     960
catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt    1020
gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc    1080
```

tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg 1140
ctcgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggaccta 1200
cgccattgtc gaaaaagaca tgcctgagga tctcaagcgc cgcttggccg actctgttca 1260
gcgaacgggc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc 1320
ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt 1380
cggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg 1440
cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cacgtcagca gctccaactg 1500
ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg 1560
ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac 1620
gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc 1680
actacgctgc tgacccagga agcgcatttg ttggacattc aggcttaccg tgcttggtta 1740
gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct 1800
tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa 1860
ctgaaagttc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg 1920
cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta 1980
cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc 2040
ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc 2100
cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg 2160
tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg 2220
aatgtacgtg tatgggcaat caacaagtcg tttcgataac cggtg 2265

<210> 28
<211> 2265
<212> DNA
<213> Artificial Sequence

<220>
<223> A modified DNA sequence

<400> 28
gagggtagag aaatcgaatg ccccttgcat caaggtcggt ttgacgtttg cacaggcaaa 60
gccctgtgcg cacccgtgac acagaacatc aaaacatatc cagtcaagat tgagaacctg 120
cgcgtaatga ttgatttgag ctaagaattt taacaggagg cacccccggc cctagagcgt 180
aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt 240
aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca 300
tgaactgaaa accatttttg cgcggaactg gcttttttctc actcatgata gcctgattcc 360
tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccggca 420
gaacgacggt tcgattcgtg cttttctgaa cgtttgccgg catcgtggca agacgctggt 480
gagcgtggaa gccggcaatg ccaaaggttt tgtttgcagc tatcacggct ggggcttcgg 540
ctccaacggt gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa 600

```
taaaaaatgt ctggggttga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta     660
cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggta     720
cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggt     780
tgtgatcaag gccaactgga aggcacccgc ggaaaacttt gtgggagatg cataccacgt     840
gggttggacg cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg     900
caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg     960
catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt    1020
gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc    1080
tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg    1140
ctcgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggaccta    1200
cgccattgtc gaaaaagaca tgcctgagga tctcaagcgc cgcttggccg actctgttca    1260
gcgaacggcc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc    1320
ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt    1380
cggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg    1440
cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cacgtcagca gctccaactg    1500
ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg    1560
ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac    1620
gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc    1680
actacgctgc tgacccagga agcgcatttg ttggacattc aggcttaccg tgcttggtta    1740
gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct    1800
tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa    1860
ctgaaagttc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg    1920
cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta    1980
cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc    2040
ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc    2100
cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg    2160
tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg    2220
aatgtacgtg tatgggcaat caacaagtcg tttcgataac cggtg                    2265
```

<210> 29

<211> 2265

<212> DNA

<213> Artificial Sequence

<220>

<223> A modified DNA sequence

<400> 29

```
gagggtagag aaatcgaatg ccccttgcat caaggtcggt ttgacgtttg cacaggcaaa      60
gccctgtgcg cacccgtgac acagaacatc aaaacatatc cagtcaagat tgagaacctg     120
```

```
cgcgtaatga ttgatttgag ctaagaattt taacaggagg caccccgggc cctagagcgt      180
aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt      240
aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca      300
tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc      360
tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccggca      420
gaacgacggt tcgattcgtg cttttctgaa cgtttgccgg catcgtggca agacgctggt      480
gagcgtggaa gccggcaatg ccaaaggttt tgtttgcagc tatcacggct ggggcttcgg      540
ctccaacggt gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa      600
taaaaaatgt ctggggttga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta      660
cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggta      720
cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggt      780
tgtgatcaag gccaactgga aggcacccgc ggaaaacttt gtgggagatg cataccacgt      840
gggttggacg cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg      900
caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg      960
catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt     1020
gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc     1080
tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg     1140
ctcgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggaccta     1200
cgccattgtc gaaaaagaca tgcctgagga tctcaagcgc cgcttggccg actctgttca     1260
gcgaacgacc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc     1320
ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt     1380
cggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg     1440
cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cacgtcagca gctccaactg     1500
ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg     1560
ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac     1620
gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc     1680
actacgctgc tgacccagga agcgcatttg ttggacattc aggcttaccg tgcttggtta     1740
gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct     1800
tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa     1860
ctgaaagttc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg     1920
cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta     1980
cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc     2040
ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc     2100
cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg     2160
tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg     2220
aatgtacgtg tatgggcaat caacaagtcg tttcgataac cggtg                     2265
```

<210> 30

<211> 2265

<212> DNA

<213> Artificial Sequence

<220>
<223> A modified DNA sequence

<400> 30
gagggtagag aaatcgaatg ccccttgcat caaggtcggt ttgacgtttg cacaggcaaa 60
gccctgtgcg cacccgtgac acagaacatc aaaacatatc cagtcaagat tgagaacctg 120
cgcgtaatga ttgatttgag ctaagaattt taacaggagg caccccgggc cctagagcgt 180
aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt 240
aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca 300
tgaactgaaa accatttttg cgcggaactg gcttttttctc actcatgata gcctgattcc 360
tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccggca 420
gaacgacggt tcgattcgtg cttttctgaa cgtttgccgg catcgtggca agacgctggt 480
gagcgtggaa gccggcaatg ccaaaggttt tgtttgcagc tatcacggct ggggcttcgg 540
ctccaacggt gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa 600
taaaaaatgt ctggggttga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta 660
cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggta 720
cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggt 780
tgtgatcaag gccaactgga aggcacccgc ggaaaacttt gtgggagatg cataccacgt 840
gggttggacg cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg 900
caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg 960
catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt 1020
gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc 1080
tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg 1140
ctcgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggaccta 1200
cgccattgtc gaaaaagaca tgcctgagga tctcaagcgc cgcttggccg actctgttca 1260
gcgaacgctc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc 1320
ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt 1380
cggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg 1440
cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cacgtcagca gctccaactg 1500
ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg 1560
ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac 1620
gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc 1680
actacgctgc tgacccagga agcgcatttg ttggacattc aggcttaccg tgcttggtta 1740
gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct 1800
tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa 1860
ctgaaagttc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg 1920
cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta 1980
cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc 2040

```
ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc    2100
cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg    2160
tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg    2220
aatgtacgtg tatgggcaat caacaagtcg tttcgataac cggtg                     2265
```

<210> 31

<211> 2265

<212> DNA

<213> Artificial Sequence

<220>

<223> A modified DNA sequence

<400> 31

```
gagggtagag aaatcgaatg ccccttgcat caaggtcggt ttgacgtttg cacaggcaaa      60
gccctgtgcg cacccgtgac acagaacatc aaaacatatc cagtcaagat tgagaacctg     120
cgcgtaatga ttgatttgag ctaagaattt taacaggagg caccccgggc cctagagcgt     180
aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt     240
aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca     300
tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc     360
tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccggca     420
gaacgacggt tcgattcgtg cttttctgaa cgtttgccgg catcgtggca agacgctggt     480
gagcgtggaa gccggcaatg ccaaaggttt tgtttgcagc tatcacggct ggggcttcgg     540
ctccaacggt gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa     600
taaaaaatgt ctggggttga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta     660
cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggta     720
cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggt     780
tgtgatcaag gccaactgga aggcacccgc ggaaaacttt gtgggagatg cataccacgt     840
gggttggacg cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg     900
caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg     960
catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt    1020
gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc    1080
tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg    1140
ctcgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggaccta    1200
cgccattgtc gaaaaagaca tgcctgagga tctcaagcgc cgcttggccg actctgttca    1260
gcgaacgatc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc    1320
ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt    1380
cggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg    1440
cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cacgtcagca gctccaactg    1500
ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg    1560
```

```
ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac     1620
gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc     1680
actacgctgc tgacccagga agcgcatttg ttggacattc aggcttaccg tgcttggtta     1740
gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct     1800
tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa     1860
ctgaaagttc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg     1920
cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta     1980
cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc     2040
ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc     2100
cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg     2160
tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg     2220
aatgtacgtg tatgggcaat caacaagtcg tttcgataac cggtg                     2265
```

<210> 32
<211> 449
<212> PRT
<213> Artificial Sequence

<220>
<223> A polypeptide encoded by SEQ ID NO:27

<400> 32

```
Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
 1               5                  10                  15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
            20                  25                  30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
        35                  40                  45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
    50                  55                  60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65                  70                  75                  80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
        115                 120                 125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
```

09843250.042601

```
His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
145                 150                 155                 160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
    210                 215                 220
Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
                245                 250                 255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
            260                 265                 270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
        275                 280                 285
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
    290                 295                 300
Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305                 310                 315                 320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
                325                 330                 335
Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Gly
            340                 345                 350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
        355                 360                 365
Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
    370                 375                 380
Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
385                 390                 395                 400
Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
                405                 410                 415
Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
            420                 425                 430
Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
        435                 440                 445
Arg
```

```
<210> 33
<211> 449
<212> PRT
<213> Artificial Sequence

<220>
<223> A polypeptide encoded by SEQ ID NO:28

<400> 33
Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
1               5                   10                  15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
            20                  25                  30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
        35                  40                  45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
    50                  55                  60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65                  70                  75                  80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
        115                 120                 125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
145                 150                 155                 160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
    210                 215                 220
Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
```

245 250 255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
260 265 270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
275 280 285
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
290 295 300
Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305 310 315 320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
325 330 335
Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Ala
340 345 350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
355 360 365
Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
370 375 380
Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
385 390 395 400
Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
405 410 415
Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
420 425 430
Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
435 440 445
Arg

<210> 34
<211> 449
<212> PRT
<213> Artificial Sequence

<220>
<223> A polypeptide encoded by SEQ ID NO:29

<400> 34
Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
1 5 10 15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys

```
            20                  25                  30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
        35                  40                  45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
    50                  55                  60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65                  70                  75                  80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
        115                 120                 125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
145                 150                 155                 160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
    210                 215                 220
Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
                245                 250                 255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
            260                 265                 270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
        275                 280                 285
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
    290                 295                 300
Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305                 310                 315                 320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
                325                 330                 335
Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Thr
```

```
                340                 345                 350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
            355                 360                 365
Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
    370                 375                 380
Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
385                 390                 395                 400
Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
                405                 410                 415
Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
            420                 425                 430
Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
        435                 440                 445
Arg


        <210> 35
        <211> 449
        <212> PRT
        <213> Artificial Sequence

        <220>
        <223> A polypeptide encoded by SEQ ID NO:30

        <400> 35
Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
 1               5                  10                  15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
            20                  25                  30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
        35                  40                  45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
    50                  55                  60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65                  70                  75                  80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
```

```
        115                 120                 125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
145                 150                 155                 160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
    210                 215                 220
Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
                245                 250                 255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
            260                 265                 270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
        275                 280                 285
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
    290                 295                 300
Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305                 310                 315                 320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
                325                 330                 335
Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Leu
            340                 345                 350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
        355                 360                 365
Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
    370                 375                 380
Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
385                 390                 395                 400
Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
                405                 410                 415
Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
            420                 425                 430
Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
```

```
                435                 440                 445

Arg


        <210> 36
        <211> 449
        <212> PRT
        <213> Artificial Sequence

        <220>
        <223> A polypeptide encoded by SEQ ID NO:31

        <400> 36
Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
1               5                   10                  15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
            20                  25                  30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
        35                  40                  45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
    50                  55                  60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65                  70                  75                  80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
        115                 120                 125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
145                 150                 155                 160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
```

```
    210                 215                 220
Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
                245                 250                 255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
            260                 265                 270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
        275                 280                 285
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
    290                 295                 300
Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305                 310                 315                 320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
                325                 330                 335
Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Ile
            340                 345                 350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
        355                 360                 365
Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
    370                 375                 380
Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
385                 390                 395                 400
Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
                405                 410                 415
Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
            420                 425                 430
Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
        435                 440                 445
Arg
```

<210> 37
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 37
ttcagcgaac ggtcgggcct gc 22

<210> 38
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 38
gccgggcctc ttgcgggata tcgtcca 27

<210> 39
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 39
gttgccattg ctgcaggcat cgtggtg 27

<210> 40
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 40
gaggcacccg cggaagcttt tgtgggagat gca 33

<210> 41
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 41
gcacccgcgg aacaatttgt gggagatgca 30

<210> 42
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 42
ccgcggaaag ctttgtggga g 21

<210> 43
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 43
ccgcggaaaa gcttgtggga gatg 24

<210> 44
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 44
cgcggaaaac gttgtgggag atg 23

<210> 45

<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 45
atattcaggt gcgcatagcg cag 23

<210> 46
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 46
ggacggatat tcagggctcc atagcgcaga cttg 34

<210> 47
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 47
gacggatatt caggtaacca tagcgcagac ttg 33

<210> 48
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 48
ggtgttttca aagtcgcgaa cccgatcgac 30

<210> 49
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 49
ctgttcagcg aaacttcggg cctgct 26

<210> 50
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 50
ctgttcagcg aaggttcggg cctgct 26

<210> 51
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 51
ctgttcagcg aagcttcggg cctgct 26

<210> 52
<211> 22
<212> DNA
<213> Artificial Sequence

09843250.042501

<220>
<223> An oligonucleotide

<400> 52
ttcagcgaac gctcgggcct gc 22

<210> 53
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 53
ggcctgctgg cttcgcggaa agcgacgaca 30

<210> 54
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 54
gaaagcgacg ccaatgacaa t 21

<210> 55
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 55
acgacaatga caattgggaa acagcttcgc 30

<210> 56

<211> 2265

<212> DNA

<213> Artificial Sequence

<220>

<223> A modified DNA sequence

<400> 56

```
gagggtagag aaatcgaatg ccccttgcat caaggtcggt ttgacgtttg cacaggcaaa     60
gccctgtgcg cacccgtgac acagaacatc aaaacatatc cagtcaagat tgagaacctg    120
cgcgtaatga ttgatttgag ctaagaattt taacaggagg caccccgggc cctagagcgt    180
aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt    240
aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca    300
tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc    360
tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccggca    420
gaacgacggt tcgattcgtg cttttctgaa cgtttgccgg catcgtggca agacgctggt    480
gagcgtggaa gccggcaatg ccaaaggttt tgtttgcagc tatcacggct ggggcttcgg    540
ctccaacggt gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa    600
taaaaaatgt ctggggttga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta    660
cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggta    720
cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggt    780
tgtgatcaag gccaactgga aggcacccgc ggaaaacttt gtgggagatg cataccacgt    840
gggttggacg cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg    900
caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg    960
catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt   1020
gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc   1080
tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg   1140
ctcgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggaccta   1200
cgccattgtc gaaaaagaca tgcctgagga tctcaagcgc cgcttggccg actctgttca   1260
gcgaacgtgg gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc   1320
ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt   1380
cggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg   1440
cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cacgtcagca gctccaactg   1500
ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg   1560
ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac   1620
gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc   1680
actacgctgc tgacccagga agcgcatttg ttggacattc aggcttaccg tgcttggtta   1740
gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct   1800
tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa   1860
ctgaaagttc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg   1920
```

cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta 1980

cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc 2040

ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc 2100

cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg 2160

tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg 2220

aatgtacgtg tatgggcaat caacaagtcg tttcgataac cggtg 2265

<210> 57

<211> 2265

<212> DNA

<213> Artificial Sequence

<220>

<223> A modified DNA sequence

<400> 57

gagggtagag aaatcgaatg ccccttgcat caaggtcggt ttgacgtttg cacaggcaaa 60

gccctgtgcg cacccgtgac acagaacatc aaaacatatc cagtcaagat tgagaacctg 120

cgcgtaatga ttgatttgag ctaagaattt taacaggagg caccccgggc cctagagcgt 180

aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt 240

aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca 300

tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc 360

tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccggca 420

gaacgacggt tcgattcgtg cttttctgaa cgtttgccgg catcgtggca agacgctggt 480

gagcgtggaa gccggcaatg ccaaaggttt tgtttgcagc tatcacggct ggggcttcgg 540

ctccaacggt gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa 600

taaaaaatgt ctggggttga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta 660

cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggta 720

cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggt 780

tgtgatcaag gccaactgga aggcacccgc ggaaaacttt gtgggagatg cataccacgt 840

gggttggacg cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg 900

caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg 960

catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt 1020

gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc 1080

tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg 1140

ctcgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggaccta 1200

cgccattgtc gaaaaagaca tgcctgagga tctcaagcgc cgcttggccg actctgttca 1260

gcgaacggtc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc 1320

ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt 1380

cggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg 1440

```
cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cacgtcagca gctccaactg    1500
ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg    1560
ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac    1620
gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc    1680
actacgctgc tgacccagga agcgcatttg ttggacattc aggcttaccg tgcttggtta    1740
gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct    1800
tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa    1860
ctgaaagttc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg    1920
cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta    1980
cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc    2040
ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc    2100
cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg    2160
tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg    2220
aatgtacgtg tatgggcaat caacaagtcg tttcgataac cggtg                    2265


<210> 58
<211> 449
<212> PRT
<213> Artificial Sequence

<220>
<223> A polypeptide encoded by SEQ ID NO:56

<400> 58
Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
1               5                   10                  15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
            20                  25                  30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
        35                  40                  45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
    50                  55                  60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65                  70                  75                  80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
        115                 120                 125
```

```
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
145                 150                 155                 160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
    210                 215                 220
Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
                245                 250                 255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
            260                 265                 270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
        275                 280                 285
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
    290                 295                 300
Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305                 310                 315                 320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
                325                 330                 335
Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Trp
            340                 345                 350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
        355                 360                 365
Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
    370                 375                 380
Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
385                 390                 395                 400
Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
                405                 410                 415
Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
            420                 425                 430
Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
        435                 440                 445
```

```
Arg


      <210> 59
      <211> 449
      <212> PRT
      <213> Artificial Sequence

      <220>
      <223> A polypeptide encoded by SEQ ID NO:57

      <400> 59
Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
1               5                   10                  15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
            20                  25                  30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
        35                  40                  45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
    50                  55                  60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65                  70                  75                  80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
                85                  90                  95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
            100                 105                 110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
        115                 120                 125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
    130                 135                 140
His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
145                 150                 155                 160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
                165                 170                 175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
            180                 185                 190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
        195                 200                 205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
    210                 215                 220
```

09843250.042601

```
Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225                 230                 235                 240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
                245                 250                 255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
            260                 265                 270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
        275                 280                 285
Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
    290                 295                 300
Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
305                 310                 315                 320
Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
                325                 330                 335
Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Tyr
            340                 345                 350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
        355                 360                 365
Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
    370                 375                 380
Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
385                 390                 395                 400
Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
                405                 410                 415
Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
            420                 425                 430
Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
        435                 440                 445
Arg
```

```
<210> 60
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 60
```

gttcagcgaa cgggcgggcc tgctgg 26

<210> 61
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 61
gttcagcgaa cggccgggcc tgctgg 26

<210> 62
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 62
gttcagcgaa cgaccgggcc tgctgg 26

<210> 63
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 63
gttcagcgaa cgatcgggcc tgctgg 26

<210> 64
<211> 26
<212> DNA
<213> Artificial Sequence

09843250.042501

<220>
<223> An oligonucleotide

<400> 64
gttcagcgaa cgtgggggcc tgctgg 26

<210> 65
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> An oligonucleotide

<400> 65
ttcagcgaac gtacgggcct gctgg 25

09843250.042601